Contract No. MDA 903-92-C-0126
Item No. 0002AJ, First Year Technical Report

Technical Report-599

Structuring and Training High-Reliability Teams Year 1 Technical Report

November 1993

Prepared by:

Eileen B. Entin
Elliot E. Entin
Jean MacMillan
Daniel Serfaty

ALPHATECH, Inc.
50 Mall Road
Burlington, MA 01803-4562

The views, opinions, and findings contained in this report are those of the author(s) and should not be construed as official Agency position, policy, or decision, unless so designated by other official documentation.

Prepared for:

Dr. Dennis K. Leedom
U.S. Army Research Institute/Aviation R&D Activity/PERI-IR
Fort Rucker, AL 36362

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
23 November 1993

3. REPORT TYPE AND DATES COVERED
Interim

4. TITLE AND SUBTITLE
Structuring and Training High-Reliability Teams:
Year 1 Technical Report

5. FUNDING NUMBERS
C-MDA-903-92-C-0126

6. AUTHOR(S)
Eileen B. Entin, Elliot E. Entin,
Jean MacMillan, Daniel Serfaty

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
ALPHATECH, Inc.
50 Mall Road
Burlington, MA 01803

8. PERFORMING ORGANIZATION REPORT NUMBER
TR-599

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute/Aviation R&D Activity
Fort Rucker, AL 36362-5354

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report describes activities during the first year of a research project sponsored by the U.S. Army Research Institute (Aviation Research and Development Activity) on the role of team coordination in reliable team performance. First-year activities included a review of the team-measurement literature, the development of a measurement approach based on a theoretical framework for team coordination, the development of a data-collection instrument for rigorous quantitative measurement of team communication and coordination patterns, and testing of the approach and the instrument based on observation of videotapes of helicopter flightcrews during simulated operations. The measures developed were found to have high inter-rater reliability as well as reasonable validity in measuring teamwork processes. Results are consistent with the hypothesis that higher-performance teams adapt to increases in workload by relying more heavily on implicit coordination—anticipating each other's needs without explicit requests or commands—as measured by the ratio of transfers to requests in the team's communications.

14. SUBJECT TERMS
Team performance, crew coordination

15. NUMBER OF PAGES
106

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

REF 8

STRUCTURING AND TRAINING HIGH-RELIABILITY TEAMS:
YEAR 1 TECHNICAL REPORT

EXECUTIVE SUMMARY

Research Requirement

The performance of teams working together to control complex systems has become increasingly important for a variety of military and civilian applications, and there is growing interest in identifying and understanding the factors that affect the reliability of such teams. There is a need to identify and measure the communication behaviors associated with effective team coordination in order to develop methods for structuring and training reliable teams. Coordination and communication measures are required in order to detect and quantify changes in the communication strategies that superior crews adopt as a result of crew-coordination training.

Procedure

First-year activities included a review of the team-measurement literature, the development of a measurement approach based on a theoretical framework for team coordination, the development of a data-collection instrument for quantitative measurement of team communication and coordination patterns, and testing of the approach and the instrument based on observation of videotapes of helicopter flightcrews during simulated flight.

Findings

The measures developed were found to have high inter-rater reliability as well as reasonable validity in measuring teamwork processes, as indicated by their sensitivity to the effects of crew-coordination training, their correlation with other validated teamwork measures, their correlation with mission-performance measures, and their congruence with a theory of crew coordination. Results are consistent with the hypothesis that higher-performance teams adapt to increases in workload by relying more heavily on implicit coordination—anticipating each other's needs without explicit requests or commands—as measured by the ratio of transfers to requests in the team's communications.

Utilization of Findings

The measurement approach and the instruments developed will be used in the second year of the project to study the effects of battle rostering and crew-coordination training on crew performance.

STRUCTURING AND TRAINING HIGH-RELIABILITY TEAMS:
YEAR 1 TECHNICAL REPORT

CONTENTS

CONTENTS

LIST OF TABLES

CONTENTS

LIST OF FIGURES

STRUCTURING AND TRAINING HIGH-RELIABILITY TEAMS: YEAR 1 TECHNICAL REPORT

Introduction

ALPHATECH, Inc., under contract to the Army Research Institute Aviation Research & Development Activity (ARIARDA) is conducting research on structuring and training high-reliability teams. This report documents the activities performed during the first year of this project. These activities included a literature review of team measures leading to the development of a recommended measurement approach, the design of instruments to be used in a crew-coordination experiment planned for the fall of 1993, visits to two Army simulation facilities used for team training, and the observation of videotapes of helicopter flightcrews during simulated flight.

Background

The performance of teams of operators controlling complex systems has become the focus of a growing research effort in recent years. There is increasing interest in identifying and understanding the factors that affect team performance and team errors in both military and civilian applications. The goal of this ARIARDA-sponsored research on structuring and training high-reliability teams is to provide insight into the development of such teams by examining how teams can best be structured and trained to support the flexible, adaptive behavior that has been observed to produce highly reliable team performance in real-world environments (LaPorte & Consolini, 1988; Pfeiffer, 1989; Reason, 1990). The effort has focused on the role of team-coordination strategies in producing reliable team performance.

First Year Activities

Progress was made during the first year of the project on four tasks:

- Task I-A. Classification of team measures. We reviewed selected literature on alternative methods for measuring team coordination and team performance in order to identify promising measures for planned research. Based on this review, and on a theory of team coordination, we developed a recommended measurement approach.

- Task I-B. Instrument development for battle rostering crew-coordination experiment. Based on a theory of crew coordination, we designed questionnaires to be used as part of an experiment examining the effects of battle rostering on crew coordination.

- Task II-A. On-site observation of team coordination and team performance. We visited two simulation facilities

to observe team behavior and team performance, to discuss team-coordination issues, and to evaluate the feasibility of conducting future experiments at those facilities.

- Task II-B. Observation and analysis of Fort Campbell videotapes. We were able to obtain and use videotapes and data collected at a helicopter flight-simulation testbed facility to test our measures, our measurement approach, and our theories about crew coordination.

Organization of this Report

The remainder of this report is divided into three major sections. The first section provides a high-level overview and summary description of first-year project activities for each of the four tasks listed above. The second and third sections provide a more in-depth treatment of the two areas in which the majority of the effort was focused: the review and classification of team measures, and the observation and analysis of communication and coordination patterns based on videotapes from the Fort Campbell Testbed.

Summary of Year 1 Progress

This section summarizes activities and accomplishments in each of the four major areas of effort during the first year of the project. Note that more detail on the first activity, "Classification of Team Measures," and the fourth activity, "Observation and Analysis of Fort Campbell Tapes," is provided in subsequent sections.

Summary of Task I-A: Classification of Team Measures

The objective of Task I was to review recent research on team performance in order to develop a set of measurement instruments that are diagnostic of team errors. A subgoal was to classify existing team measures into a taxonomy in order to identify overlaps and areas where additional measures are needed. This subsection summarizes activities and recommendations from Task I. The literature review, the comparison of alternative measurement dimensions, and the recommended approach for studying crew coordination are discussed in more detail in the second section of this report, "Classification of Team Measures."

We reviewed the following team-measurement instruments and approaches in Task I:

- the Aircrew Coordination Evaluation (ACE),
- the Critical Team Behaviors Form (CTBF),
- the AAW Teamwork Observation Measure (ATOM),
- the Teamwork Needs Assessment Scale (TNAS),

- measures used in Distributed Dynamic Decisionmaking (DDD) research,
- the Cockpit Management Attitude Questionnaire (CMAQ),
- measures based on semantic analysis of communication content, and
- measures based on the Aircrew Training Manual (ATM).

A theoretical framework or model was needed to guide the selection and development of crew-coordination measures. We proposed such a framework and, based on that framework, identified the following dimensions for the required measures: 1) crew workload, 2) crew processes, including both taskwork and teamwork, and 3) crew performance. Crew coordination falls into the area of teamwork, and our review and measure-development effort focused specifically on process measures of teamwork.

We identified two major approaches to teamwork-process measurement. The first relies on ratings by domain experts of the dimensions of crew behavior and teamwork that they believe to be important for effective crew performance. The CTBF, the ATOM, and the ACE use this approach. The second approach analyzes crew communications and crew actions in detail and relies on event counts rather than ratings by experts. The DDD measures and the communication-content measures are based on this approach. We recommend the use of both measurement approaches for future crew-coordination research. For flightcrew experiments, we recommend use of the ACE to obtain expert ratings of crew coordination, and the development of communications-analysis instruments, based on DDD and communications-content measures, to obtain quantitative data on communication patterns.

Figure 1 shows the methods and measures recommended for obtaining the data suggested by the theoretical framework for flightcrew coordination. We recommend the use of existing crew-workload measurement approaches and the use of the previously developed ATM-based performance measures for flightcrews as well as the use of mission-specific performance indices. For coordination and teamwork process measurement, we recommend both the use of existing teamwork-process measures such as the ACE and the CMAQ, and the development of new quantitative communications-analysis measures. During Year One we developed a communication-analysis data-collection instrument and methodology to provide these measures and, under Task II-B, were able to test the instrument using videotapes of flightcrews during simulated flight.

<u>Summary of Task I-B: Instrument Development for Battle Rostering Crew-Coordination Experiment</u>

During Year One we developed questionnaires to explicitly test the mental-model assumptions underlying our theory of team

MEASURE / METHOD	WORKLOAD	CREW PROCESSES		PERFORMANCE (OUTCOME)
		TASKWORK	TEAMWORK	
SELF-REPORT (CREWS)	TLX		CMAQ	
AUTOMATED (SIMULATOR)		MISSION-SPECIFIC INDICES		MISSION-SPECIFIC INDICES
ON-SITE/OFF-SITE OBSERVATIONS (OBSERVERS)			COMMUNICATION ANALYSIS	
ON-SITE OBSERVATIONS (DOMAIN EXPERTS)		ACE EVALUATION QUESTIONNAIRES		ATM, SUBJ. EVAL.

Figure 1. Recommended measures for flightcrew-coordination research.

coordination. We plan to collect data using the questionnaires in conjunction with an experiment sponsored by ARI on crew-coordination training and the effects of battle rostering for crews.

The premise of the research is that *congruent mental models among crew members lead to superior crew performance.* This "congruence" has at least three dimensions:

- Congruence with "truth": i.e., how close to the truth are the two crew members' assessments of the situation?
- Congruence of the two situational mental models: i.e., how close to each other are the two crew members' assessments of the situation?
- Congruence of the mutual mental models: i.e., how well does one crew member anticipate/predict/understand the actions and information needs of the other?

Deficiencies in any of these three dimensions may be evident before, during, or after the mission. We expect that different deficiencies will be evident at different times. Therefore we must collect data and assess mental-model congruence at several points in time.

An associated hypothesis for the research is that battle rostering for crews may produce an exaggerated belief among crew members that they can anticipate and predict the actions and needs

of the other crew members. This may lead to a variety of *team errors that differ from the errors made by a crew that is not battle rostered.*

The goals of our research effort are to assess the extent to which crew members hold shared congruent mental models of a mission and to assess whether congruent mental models of crew members will lead to higher performance. We plan to gather two kinds of evidence to test the congruence of crew member's mental models and the relationship of that congruence to performance, as shown in Figure 2. We will gather cognitive evidence through questionnaires, and behavioral evidence primarily from videotape observation.

COGNITIVE EVIDENCE		BEHAVIORAL EVIDENCE
Measures: Match/mismatch of pieces of mental model	predictive relationships ⟷	Measures: Team communication patterns & crew coordination measures
Where: pre & post mission		Where: Real-time, pre & post mission
SOURCES:		SOURCES:
1. Crew questionnaire during prebrief 2. Crew questionnaire after mission 3. Instructor-pilot questionnaire after mission		1. Direct observation of crews 2. Videotapes (analyses)

Figure 2. Data collection for battle rostering experiment.

Instruments

We developed three questionnaires to assess crew members' mental-model congruence during the battle rostering experiment. These measurement instruments are compatible with the measurement approach outlined in Figure 1. Two questionnaires were designed specifically for the subjects (crews), and the third for the instructor pilots (IPs).

The Crew Member-Mission Questionnaire is completed independently by each crew member prior to the pre-training scenario and again prior to the post-training scenario. Four items elicit elements of the crew members' mental models of critical aspects of the mission. Two other items address the issues of mutual confidence and perception of mutual confidence among the crew members. The purpose of the questions is to assess the congruence of the situational and mutual mental models of crew members prior to the mission.

The second instrument is the Crew Member Post-Mission Questionnaire. Each crew member is asked to complete the instrument independently after the pre-training scenario and again after the post-training scenario. These items examine the extent to which each crew member was able to anticipate (i.e., predict) the actions and decisions of the other. The goal is to assess the congruence of situational and mutual mental models of crew members subsequent to mission performance.

The third questionnaire, the IP Post-Mission Questionnaire, is completed by the IPs, who are domain experts. The questionnaire is completed for each crew following the pre-training scenario and again following the post-training scenario. The 15 items of this instrument are constructed around six dimensions of teamwork and various aspects of crew members' mental models, including the extent to which each crew member predicts the other's actions and decisions, the extent to which each crew member understands the actions and decisions of the other, the extent to which crew members share workload, and the extent to which crew members support one another. Each item in the questionnaire is accompanied by a seven-point scale, with behavioral anchors that describe the two ends of the scale. The three questionnaires developed for the experiment are shown in Appendix A.

Summary of Task II-A: On-site Observation

The objective of Task II was to observe teams in naturalistic environments in order to identify coordination issues, assess the feasibility of planned measurement approaches, and assess the feasibility of conducting crew-coordination experiments in those environments. We visited two sites under Task II that we believed might be suitable for systematic observation and experiments investigating team-coordination strategies. One potential site was the Simulation Network (SIMNET) facility at Fort Knox, KY. The second was the Army's Combat Training Center (CTC) data archive at the Presidio of Monterey, CA.

SIMNET Facility

We visited the SIMNET-T (training) facility at Fort Knox to evaluate the feasibility of using this facility for our research on team communication and coordination. We toured the facility with Dr. William Burnside (from ARI Fort Knox), observed a training exercise, and conducted an interview with one of the officers involved in simulation training.

SIMNET supports 58 individual M3 armored personnel carrier and M1 tank simulators. The tank crews communicate via an intercom system within the tank and via radio with other tanks. Observers can view an ongoing scenario from the "stealth" observation facility, which enables observers to see the battlefield from different vantage points. Three large screens show the scenario as it would be seen from above (as if in a

helicopter), adjacent to (beside, in front of, or behind) any vehicle, or inside any vehicle. Another display portrays an electronic topological map of the whole area with icons showing the positions of friendly and enemy forces. As the forces move, the map is updated. An observer can place time markers in the simulation so that any designated segment of a scenario can be retrieved and displayed during the after-action review (AAR).

During our visit we observed a training exercise. The stealth station provided a good vantage point for observing the action. All movement could be followed on the electronic map, and the details displayed on the large screen could easily be seen. Monitoring the communications was considerably more difficult. Unless one was well versed in armor tactics and Army acronyms, understanding the communications was almost impossible.

As part of our visit we interviewed an armor company commander who is involved in simulation training. The goal of this interview was to explore the coordination issues that: 1) exist among crew members in a single vehicle, 2) exist within a platoon of tanks (i.e., four vehicles), and 3) exist within a company of tanks (i.e., three platoons).

Coordination is most important between the tank commander and his driver. The driver must learn how to follow (and anticipate) the tank commander's instruction, while exhibiting good terrain driving and tank orientation. The coordination between the tank commander and gunner is also important. The gunner assists the tank commander in scanning for the enemy and should anticipate the commander's target selections.

Crews must master all the technology in a modern armored vehicle, but they must also know how to operate the systems in degraded or manual mode, which demands higher coordination among crew members. Another challenge to crew coordination is turnover. An average crew stays together for two to four months and then one or two crew members are replaced by new personnel. This practice may be disruptive to a crew's operating behavior.

Coordination issues in a platoon include knowing own-force location (particularly at night or traveling with the hatch closed), staying in formation, establishing the right formation, communicating effectively without using the radio (i.e., flags, hand signals, lights), exhibiting good net discipline when using the radio, and coordinating fire distribution (to insure that everyone is not shooting at the same target). The most difficult coordination issue for the company commander and his executive officer is the logistics of fuel supply. Coordination with Army infantry, air, and artillery is usually good as long as the radios work, but interservice links and coordination may be a problem.

On the basis of our visit, we concluded that while SIMNET-T is an impressive distributed-training facility, it does not at this time provide all of the capabilities needed for our planned

research on team coordination. A number of the circumstances that require the most intensive coordination are not captured in the current version of the simulation. For example, the problems encountered in operating with systems in a degraded or manual mode, traveling at night, or communicating without the radio are not represented in the simulation. The observational capability of the stealth position is focused on platoon or company maneuvers, and provides the information required for training platoon- and company-sized units. Communications among crew members within a tank are not captured at all, and communications between tank crews are poorly recorded. We do not see an easy way to overcome these problems. We could place audio recorders in one or more of the tank modules to record ambient crew conversations, but it is not clear that inter-tank conversations would be captured. Nor is it clear how conversations within or between tanks could be synchronized in time. These problems are not insurmountable, but the solutions may demand more resources and time than are available, and perhaps more than the effort merits.

The CTC Data Archive

The CTC data archive provides a facility for storing data from the National Training Center (NTC), the Joint Readiness Training Center (JRTC), and the Combined Arms Military Training Center (CMTC). The archives include recordings of radio nets, videotapes of unit AARs, operation orders and graphics, observer/controller assessments of unit performance, data on task performance, and digital battle replay tapes. A portion of the data in the archives has been entered into a relational data base, but the bulk of it is in the form of audio tapes (almost a million hours), videotapes, overlays, and paper documents.

We discussed the possibility of using the data archives for coordination research with Dr. James Banks, a senior member of the Army Research Institute (ARI) staff at the Presidio of Monterey. The data base includes some performance metrics for both the Blue and the Red forces (for example, targets, losses, and rounds expended). Although there are occasional gaps in the data, the information in the data base provides a good picture of how the Blue units performed against the opposing forces (OPFOR). On the other hand, the data base does not typically include information about the OPFOR plan because that information is rarely supplied to the archive. Some of the evaluations made by observer/controllers can be accessed in the data base, but the remainder—the larger portion—must be retrieved from the original recording forms.

Despite the large quantity of information in the data base, there is little information to support a team coordination-strategy analysis without a significant effort in data reduction. It would be necessary to use either the audio tapes or the AAR videotapes to obtain the information needed for an analysis of team-coordination strategies. In order to link measures of team coordination to measures of team performance, we need at least

some measures of a unit's performance. This information is not included with the data that arrives from the various Centers. ARI has been developing global measures to rate the performance of units in exercises, but these have been computed only for selected tapes. Dr. Banks felt that we would need the help of a subject-matter expert and some documentation of the exercise plans (in the archive) to help decipher the recorded communication traffic and to understand the context in which the communications occurred.

Based on the information we received, we concluded that there are a number of problems that would require solutions before we could use the data base for our investigation of team coordination. Resolution of these problems would require a considerable investment of time and resources.

Conclusions and Recommendations from Task II-A

Based on our site visits to the SIMNET facility and the NTC data archive, we reached the following conclusions and recommendations:

- The two facilities we visited do not readily lend themselves to our planned research on team coordination.
- The modifications and enhancements that would be required to test our hypotheses of interest are beyond the time and resource budget of this effort.
- We recommend that other avenues be explored for the conduct of our observations and experiments.

Fortunately, we were able to identify an alternative source of data for testing our data-collection approach and our theoretical framework for crew coordination. This alternative involved the observation of videotapes of helicopter flightcrews during simulated operations, as discussed in the next subsection.

Summary of Task II-B: Observation of Fort Campbell Videotapes

One of the major objectives of Task II was to perform initial testing of our theoretical framework for team coordination and the measures based on that framework through the observation of teams in a naturalistic environment. We had an opportunity to perform this testing through the use of materials from a coordination-training experiment conducted by the Army Research Institute Aviation Research and Development Activity (ARIARDA) at Fort Campbell, Kentucky. These materials included videotapes of two-person UH-60 helicopter flightcrews during a simulated scenario, teamwork process data collected using the ACE and CMAQ instruments, and crew-performance data based on Aircrew Training Manual (ATM) measures and mission-specific indices. We used the videotapes to test the reliability of our communication-analysis measurement instrument and methodology, and the accompanying data to assess the validity of the communication and coordination

measures produced from the videotape. The details of the instrument, the methodology, and the analysis results are presented in the third section of this report, "Observation and Analysis of Fort Campbell Tapes."

We obtained videotapes of 16 crews during two simulated-flight scenario sessions. One session was conducted before the crews received coordination training, and the other after the crews received training. A team of two raters observed the videotapes using a communication-analysis data-collection form to record counts of the number of communications by type (transfers or requests) and by function (information exchange, actions or tasks, or planning or problem solving) during routine and crisis conditions throughout the scenarios. The data from one team was used for rater training and calibration, and the raters coded five tapes in common in order to produce data for an assessment of inter-rate reliability.

We found that the inter-rater reliability for our instrument and methodology, as measured by coefficient alpha, was quite high. The mean coefficient alpha for 50 measures over five tapes was .82 out of a maximum possible value of 1.0 (total agreement). We conclude that the raters, who were not domain experts in helicopter operations, were able to reliably categorize and record the number of crew communications by type and function.

We computed three major types of communication and coordination measures from the data collected. First, we computed anticipation ratios, measured as the number of information transfers divided by the number of information requests. The anticipation ratio is a measure of the extent to which each crew member anticipates the needs of the other. We also computed communication rates, defined as the volume of communications per minute, and the proportion of communications by function, i.e., the proportion of communication that involved information exchange, actions or tasks, or planning or problem solving.

We evaluated the validity of these communication and coordination measures from several different perspectives. First, we evaluated the sensitivity of the measures to the effects of crew-coordination training. Next, we evaluated the relationship of the measures to other measures of teamwork from the ACE and the CMAQ, as well as the relationship of the measures to crew-performance measures from the ATM and other mission-specific indices. Finally, we assessed the congruence of the results with our theoretical framework for crew coordination to see if the framework could explain the findings.

Overall, we conclude that the communication and coordination measures showed reasonable validity from all the perspectives tested. We found that the training provided to crews increased anticipatory behavior as measured by the anticipation ratio, as well as increasing communication rates. We also found that training changed communication functions during crisis conditions.

Before training, crews showed no differences in their communication functions during routine and crisis situations. After training, they showed proportionately more planning and problem-solving behavior and less information exchange during routine conditions. In crisis conditions, there was less communication about actions and tasks, and more exchange of information.

The communication and coordination measures based on the videotape observation were correlated with the basic dimensions of teamwork measured by the ACE and with the attitudes measured by the CMAQ. Higher anticipation ratios, more communication (after training), and a higher proportion of communication involving planning and problem solving were positively correlated with a number of ACE dimensions that indicate better teamwork, but there was no one-to-one mapping between our measures and the ACE dimensions. We also found that anticipation ratios and higher communication rates (after training) were related to a number of overall mission-performance measures. We conclude that our measures are related to better teamwork and to better crew performance as measured by other instruments.

The data available from the testbed were not sufficient to fully test the crew-coordination theory, but the results support those parts of the theory that could be tested. The theory suggests that crews have mutual mental models of other crew members as well as a shared mental model of the situation confronting them. These mental models allow the crew members to anticipate the needs of the other crew members and the requirements of the mission without the need for explicit requests and commands. This implicit coordination, based on mental models, is most important during high-workload conditions because it allows the well-trained crew to maintain an acceptable level of performance under high workload.

The anticipatory behavior observed on the videotapes supports the theory of implicit coordination. We found that coordination training increased the anticipation ratio, and that a higher anticipation ratio after training was associated with better crew performance. We had no explicit measure of workload in the available data but, assuming that workload was higher in the crisis conditions as defined in the scenarios, we found that anticipation ratios were higher in crisis than in routine conditions as predicted by the theory. This difference was found both before and after training.

The remainder of this report provides more detail on the two most resource-intensive first-year tasks—the literature review and measure-development efforts in Task I-A, and the observation and analysis of communication patterns based on the videotapes from the Fort Campbell testbed in Task II-B.

Detailed Report on Task I-A:
Classification of Team Measures

The first major task for the project was to review and assess the literature on measuring team coordination and team performance, and to recommend measures for exploring the role of team coordination in high-reliability teams. This section reports the results of that task in detail. We begin with a discussion of the motivation and objectives of the task, followed by a selective review of the team-measurement literature. We then propose a theory of team coordination, examine the implications of that theory for the dimensions of coordination that must be measured, and evaluate the adequacy and relevance of existing approaches for providing those measures. We conclude with a discussion of the selection and development of specific measures for planned flightcrew-coordination experiments, based on the theory of team coordination and on previous instrument-development efforts.

Motivation

There has been a renaissance of team performance research in the last ten years, motivated by the advent of distributed systems and decentralized operations in both the civilian and military sectors (Barrett, 1993; Swezey & Salas, 1992). While the small-group research performed by social psychologists in the 1950s and the 1960s focused on restricted face-to-face problem-solving situations, this new trend concentrates on distributed decision making, in which each team member has specialized knowledge and expertise and operates in a fast-paced, coordination-intensive environment. For example, distributed tactical decisionmaking for naval and land-based military operations is a domain that has generated a wealth of results in team performance research (Vaughan, 1990).

On the other hand, the development of measures in human factors research is far from being an exact science (Meister & Enderwick, 1992). This situation is even more acute in the team-performance measurement domain. There is a diversity of independent efforts, each creating its own set of measures specific to particular tasks. The absence of a general theory of team behavior, and the resulting lack of generalizable measures, hinders systematic comparisons across studies. We need a classification of existing team measures and attributes in order to recognize overlaps between the various approaches and to identify areas where additional measures are needed. Our initial task in this project was to take a first step toward such a classification, focusing specifically on team coordination. This section reports the results of that task.

Objective

The objective of Task I-A was to review a selection of recent research efforts in team performance and crew coordination in an attempt to classify existing measures into a taxonomy of team

measures. The classification focuses on *teamwork* measures (behaviors related to team-member interactions, such as communication and coordination) rather than on taskwork measures (behaviors related to tasks performed by the team as a single unit, such as external information seeking). The goal of this review is to provide team-performance and crew-coordination researchers with generic sets of team dimensions from which to derive appropriate measures that can be tailored for a particular application.

Literature Review

Several measurement approaches have been used to study team behavior and team performance over the past decade (Foushee & Helmreich, 1988). The detailed measures used in this previous body of research were shaped by the nature of the tasks being studied and the focus of the research. In this subsection we provide a high-level overview of the measures developed and their applications.

The Aircrew Coordination Evaluation (ACE)

The Aircrew Coordination Evaluation (ACE) checklist was developed by Simon (1991) to measure team or crew coordination. The ACE was developed to support the training of Army aviation crews by providing an assessment of the quality of the crew's coordination. The ACE is administered by trained evaluators who are subject matter experts, usually IPs. The evaluators rate the crew being observed using 19 different measures, each rated on a seven-part scale. In subsequent studies these 19 measures were reduced to a set of 13 "basic qualities."

The 13 ACE basic-quality measures deal with crew coordination behaviors, including factors such as "decision communicated or acknowledged" and "workload effectively distributed/ redistributed." These crew-coordination factors are rated from "very poor" to "superior." Three of the measures deal with overall mission performance and workload, rated from "very low" to "very high." Two of the measures deal with non-standard situations or behaviors, with the crew's emergency management and conflict resolution rated from "very poor" to "superior."

Recent use of the ACE (Simon, 1991) has shown it to be quite reliable. The ACE was also found to be strongly related to mission-performance measures. Factor analysis shows that 72 percent of the variance in crew scores can be accounted for by grouping the ACE measures into three factors: communication and group climate, workload and performance management, and cross monitoring by crew members.

Critical Team Behaviors Form (CTBF)

The Critical Team Behaviors Form (CTBF) was developed by Oser, McCallum, Salas, and Morgan (1989) as a tool for

understanding what constitutes teamwork in order to develop training strategies that enhance team performance. The CTBF was developed employing a critical incident approach. Oser et al. interviewed instructors at the Naval Gunfire Support School, reviewed the School's training materials, and observed the School's training exercises in order to identify behaviors linked to team success or failure during training. The behaviors identified were categorized into seven behavioral dimensions:

- communication,
- cooperation,
- team spirit and morale,
- giving suggestions and criticism,
- accepting suggestions and criticism,
- coordination, and
- adaptability.

A checklist was developed to assess the occurrence of 68 behavioral items in these seven categories. One-half of the items are rated as "effective" (i.e., behaviors common to successful teams) and one-half are rated as "ineffective" (i.e., behaviors common to unsuccessful teams). Instructors mark the behaviors that they observe on the CTBF during a training session. The CTBF does not require evaluators to make a judgment on a good-to-bad quality scale, but rather to note the frequency with which each specified behavior is observed (from "rarely" to "consistently").

Teams with high performance levels are expected to demonstrate many of the effective behaviors and few of the ineffective behaviors on the CTBF, while the opposite pattern is expected for teams with low performance levels. Oser et al. (1989) report empirical results indicating that the CTBF can discriminate between teams that show high performance and those that show poor performance.

AAW Teamwork Observation Measure (ATOM)

The AAW Team Observation Measure (ATOM) as reported by Dwyer (1992) is an adaptation of the CTBF designed to assess team dynamics and performance in anti-air warfare (AAW) activities in a naval environment. The ATOM measures seven components or attributes of teamwork based on, but not identical to, the dimensions used in the CTBF:

- team leadership,
- team orientation,

- communication,
- monitoring,
- feedback,
- back-up behavior, and
- coordination.

The rater judges the level of skill of the team on each of these dimensions using a five-point scale from "hardly any skill" to "complete skill," with "adequate skill" as the mid-point on the scale. The rater is provided with behavioral examples, based on the CTBF behaviors but specific to the AAW environment, for the low, mid-point, and high ratings on the scale for each of the seven dimensions.

The Teamwork Needs Assessment Scale (TNAS)

Stout, Cannon-Bowers, Morgan, and Salas (1989) have developed the Teamwork Needs Assessment Scale (TNAS) to measure the degree of teamwork *required* in specific training situations. The ultimate purpose of this assessment is to focus team-training resources on situations where they are most needed, i.e., where teamwork is most critical.

The TNAS may be filled out by team members or by an observer. It contains 35 items derived from a search of the team training/performance literature. The TNAS is organized into two sections—behavioral items and situational items—with each item rated on a seven-point scale. For the behavioral items the raters judge the extent to which each behavior is required for success in the situation. Behavioral items include, for example, "Members learn to assist other members who become overloaded," and "Members pace their activities to fit the needs of the group." For each item a high rating is believed to indicate a high need for teamwork, while a low rating indicates a low need. The situational items require that the rater judge the extent to which each item applies to the situation. An example situational item is "Tasks can be reallocated among members during performance." High ratings on the situational items are also believed to indicate a high requirement for teamwork.

Measures Used in Distributed Dynamic Decisionmaking (DDD) Research

Kleinman and his colleagues (e.g., Kleinman, Pattipati, Luh, & Serfaty, 1992; Kleinman & Serfaty, 1989) have studied team behavior and performance using the Distributed Dynamic Decisionmaking (DDD) paradigm—an abstract laboratory environment that recreates many of the task demands of distributed command and control decision making. An output of this research has been a set of 120 team measures that have been validated in numerous empirical studies. The types of measures that have been collected

and analyzed in the various experiments fall roughly into four (not necessarily disjoint) categories.

Performance measures. These measures capture team-level outcomes. They include team reward earned, team accuracy, and the timeliness and effectiveness of the team in processing information and/or tasks. These measures are often broken down into sub-measures by individual team member, by task type, or by time in which actions were taken.

Decision strategy measures. These measures describe the mechanisms by which the team attained its performance. These process measures capture the team as a "black box" and do not capture team-interaction activities. They include items such as: information-seeking strategies among the individual team members, team resource utilization, team decision-making strategies, and team failure to perform certain tasks.

Coordination measures. These measures describe the means by which team strategy was effected. They include processes inside the black box. These measures also include communications usage and describe total communication patterns among the team members. These can be broken down further into categories according to whether the communication involved information, resources, or actions, and also by whether the communication was a request or a transfer (e.g., solicited versus unsolicited communications). Other measures are indicators of conflicts in decision making caused by ineffective coordination. These include the number of times that two or more team members processed the same task (thereby wasting time and resources), the number of missed tasks under joint responsibility, timing errors in resource or information transfer, etc. Dominant measures in this category are anticipation ratios that measure the ability of the team members to use implicit-coordination strategies by anticipating their teammates' needs and providing information and resources voluntarily (Serfaty, Entin, & Volpe, 1993a.)

Workload measures. These are subjective measures obtained using the SWAT technique (Reid, Shingledecker, Nygren, & Eggemeir, 1981) that classifies workload along the three dimensions of time load, mental effort, and psychological stress. Kleinman and Serfaty's studies have pioneered the use of SWAT to elaborate and assess *team workload* in addition to individual workload. Administration of the SWAT is accomplished in two phases: a calibration phase and an assessment phase. During the first phase, usually occurring after training or practice trials, each subject is asked to sort a pack of 27 cards derived from crossing three levels of time load × three levels of mental effort × three levels of psychological stress into an ascending order of workload. Subsequently the same process is performed by the *team as a whole* in a consensus-reaching exercise. The resulting data are subjected to conjoint measurement analysis to yield a weighted workload scale for each subject, as well as for the team as a whole. During the second phase (data collection), each subject

rates his or her immediate task experience on the three dimensions on a three-point scale. Subsequently, the team members meet to provide consensus-driven ratings as a team. The three ratings are then weighted and combined according to the individual subject's scale—and to the team's scale—resulting in normalized workload scores. For purposes of comparison, the NASA Task Load Index (TLX) has recently been used in DDD experiments (Kalisetty, Kleinman, Serfaty, & Entin, 1993.) The TLX is a well-documented subjective workload assessment instrument developed at NASA (Hart & Staveland, 1988). From a practical viewpoint, the main advantage of the TLX over the SWAT is that it requires less-extensive calibration prior to its use.

Cockpit Management Attitude Questionnaire (CMAQ)

The Cockpit Management Attitude Questionnaire (CMAQ) was developed by Helmreich (1984) and Helmreich and Wilhelm (1986, 1987) to assess the attitude of commercial aviators regarding team training. The CMAQ is based on the hypothesis that attitudes in the cockpit are linked to performance, and Helmreich and Wilhelm (1986) demonstrated such a link. The CMAQ has three primary dimensions that form the subscales of Coordination and Communication, Recognition of Stressor Effects, and Command Responsibility.

Recently the CMAQ has been modified for use with military crews (Geis, 1987; Simon, 1991; Wagner, Simon, & Leedom, 1990). The Army CMAQ was developed as a questionnaire asking aviators to rate the extent of their agreement or disagreement with 46 statements concerning attitudes toward aircrew coordination, using a seven-point scale anchored at one end by "strongly disagree" and at the other end by "strongly agree." The Army CMAQ includes items such as "The pilot-in-command should use his crew to help him maintain situation awareness" and "Crewmembers should be aware of the other crewmembers' workload."

Measures Based on Semantic Analysis of Communication Content

Orasanu (1990) analyzed the content of all communications among crew members during a simulated flight task in order to identify differences in the cognitive and communications strategies used by high-performance and low-performance crews. Orasanu measured the frequency of utterances dealing with problem solving, resource management, and standard operating procedures. Each of these categories was further divided into subcategories such as recognizing problems, stating goals and subgoals, gathering information, coordinating and managing time, and sharing information. Orasanu found a number of differences in the frequency with which different types of utterances were made by crew members in high-performance versus in low-performance crews during the normal (low-workload) and abnormal (high-workload) phases of a flight scenario. Orasanu and Fischer (1992) extended this work by focusing on the metacognitive/problem-solving dimension. Six types of utterances were selected for analysis of

cockpit communications between two- and three-person crews: goals, planning, predictions, explanations, task allocations, and direct commands. All of these categories were coded by personnel knowledgeable about aircraft systems and procedures.

Measures Based on Aircrew Training Manual (ATM)

Simon (1991) developed aircrew performance measures based on the tasks delineated in a helicopter aircrew training manual (ATM). Training-manual tasks include navigation, threat prosecution, evasive maneuvers, fuel management procedures, and instrument-approach landing. These tasks were subsequently used as the basis for performance measures such as time to cross the forward line of own troops (FLOT), number of flight deviations, and percentage of time off course. Some of these mission-performance measures can be objectively scored using data captured by the flight simulator. Simon also provided IPs with a list of tasks from the ATM and asked them to score how well each task was completed by each crew using the "grades" of A, B, C, and U. Although this comprehensive set of measures is aimed at assessing the taskwork performance of a crew, it can be used as a correlate for measures of teamwork such as the ACE.

Summary of Previous Approaches

The studies discussed above show a variety of approaches to the measurement of team coordination. The ATOM measures coordination at a very high level, asking only for an overall rating of the level of skill shown by the team in communication and coordination. Both the ACE checklist and the CTBF identify observable behaviors associated with coordination at a more-detailed level, and task an observer to look for these behaviors. The ACE requires observers to rate behaviors on a directional scale, while the CTBF asks merely for a rating of "present" or "absent" for each behavior. The CMAQ takes a completely different approach, measuring crew attitude rather than crew behavior. The DDD measures seek to *quantify* the frequency and distribution of different *types* of communication and coordination behaviors, and to develop coordination indicators based on objective task performance, such as duplication of effort and anticipation ratios. Orasanu's analysis of communication content takes this quantification to an even more detailed level by measuring the frequency with which each crew member produced utterances in each of a large number of functional categories.

Theoretical Framework for Crew Coordination

A theoretical framework or model is needed to guide the selection and development of crew-coordination measures. Without such a model, there is no guidance as to what is important to measure in a crew-coordination study. The lack of theories to guide measurement is recognized as a general problem for team-performance research. In reviewing alternative approaches for the measurement of teamwork skills, Baker and Salas (1992) conclude

that "many teamwork measures are unreliable, complex, and insensitive and measure irrelevant variables," and they strongly urge the use of theory to dictate the aspects of teamwork that are to be measured and evaluated.

One cannot understand the subtleties and richness of crew-coordination strategies without establishing links between crew-coordination strategies and crew performance, as well as identifying coordination strategies that are likely to lead to errors. Furthermore, since most critical crew errors seem to happen in periods of high workload, it is also important to understand the coping mechanisms that crews use to adapt to workload while attempting to maintain their effectiveness. Therefore, a rigorous measurement tool should measure crew-communication patterns quantitatively and in enough detail to link crew coordination strategies to superior team performance, trace back the occurrence of team errors to specific coordination patterns, and assess the effects of workload on both coordination and performance.

In order to meet these goals we need a theoretical construct that will link workload, crew processes (teamwork and taskwork), and performance (outcome). Figure 3 describes such a model proposed and validated by Serfaty and his colleagues (see, for example, Serfaty, Entin, & Volpe, 1993b). It is based on the premise of *adaptation:* Superior crews cope with increases in workload through internal mechanisms of decision-strategy and coordination-strategy adaptation in an effort to keep team performance at the required level while maintaining workload at an acceptable level.

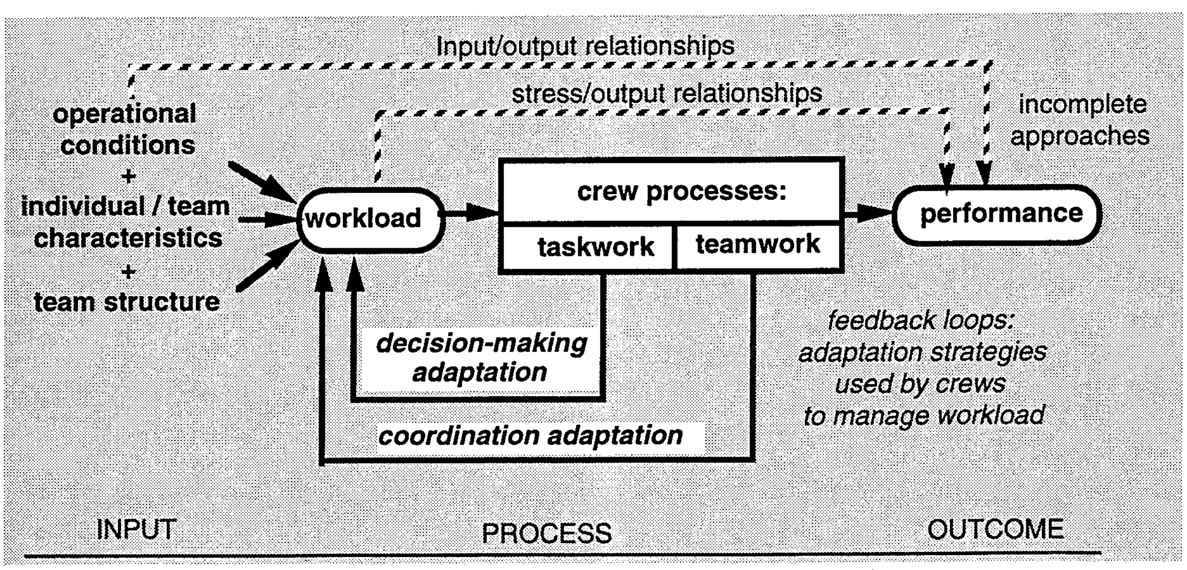

Figure 3. Theoretical framework for crew-coordination and crew-process dimensions and measures.

A key teamwork coping mechanism used by experienced crews is the shift from an explicit to an implicit coordination mode. Implicit coordination does not necessarily imply a reduction of intra-team communications. Rather, it involves a re-allocation of communication resources, a selective strategy as to the type of utterances expressed by the team leader during crisis (e.g., commands vs. explanations), a higher proportion of unsolicited information transfers (anticipation), and—depending on the context—a different mix of information granularity (raw data/processed information). In short, implicit coordination is observed not only through a reduction and redirection of communication traffic, but also through a different mix and context of messages. The implicit coordination mode involves the use of team members' mental models (Cannon-Bowers, Salas, & Converse, 1990; Orasanu, 1990) made up of two principal components: a *shared or congruent mental model* of the situation among team members, and a set of *mutual mental models* about the other team-members' states, functions, and abilities. The existence of shared mental models and the associated use of implicit coordination result in a marked improvement in team performance and a drastic reduction of team errors in high-workload conditions (Serfaty, Entin, & Deckert, 1993).

In the absence of a general theory of team behavior and crew coordination, a theoretical framework such as the one shown in Figure 3 can be used to link crew processes to team performance. The dynamic processes occurring in the cockpit during abnormal or crisis-like situations generate substantial levels of workload for the crew members. As a result, their behaviors and cognitive strategies—both individual and team-related—are strongly contingent upon the task environment. A good adaptation in coordination strategies may result in superior performance. On the other hand, a maladaptation or a lack of adaptability on the part of the crew may result in catastrophic errors.

The framework of Figure 3 suggests that three categories or types of measures are needed to understand crew coordination: 1) workload measures; 2) crew process measures, including both taskwork measures and teamwork measures; and 3) performance measures. Figure 4 shows the team-measurement approaches and instruments discussed above that provide information in each of these categories. Team workload has been measured using both the SWAT and the TLX. Team Processes have been measured using a variety of approaches, including the ACE; the CTBF and ATOM; the CMAQ; the DDD communication, coordination, and decision strategy measures; and Orasanu's communication-content measures. Performance has been measured both through evaluation by domain experts (the ATM measures) and through the analysis of simulator data on performance (ATM Measures and DDD performance measures).

The distinction between the taskwork and teamwork dimensions of team processes shown in Figures 3 and 4 is not uniformly accepted in the literature on crew coordination and team performance. These two concepts are intuitively appealing,

	TYPE OF MEASURE			
METHOD FOR DATA COLLECTION	WORKLOAD	CREW PROCESSES		PERFORMANCE (OUTCOME)
		TASKWORK	TEAMWORK	
SELF-REPORTS	SWAT TLX		CMAQ	
RATINGS AND COUNTS -Obervers -Domain Experts Simulator Data		ACE DDD Decision-Strategy Measures	ACE CTBF/ATOM DDD Communication and Coordination Measures Orasanu Communication Measures	ATM Measures Mission-performance measures DDD Performance Measures

Figure 4. Team-measurement approaches and instruments by type of measure.

however, and are practical if one wishes to distinguish between the team viewed as a single cognitive unit (taskwork) and the team viewed as a network of interactions among decisionmakers or crew members (teamwork). For example, the team total information-seeking activity from external sources is classified as a taskwork dimension, while the information exchanges among the team members are classified as a teamwork dimension.

Selection of Measures for the Analysis of Flightcrew Coordination

Based on the theoretical framework for crew coordination illustrated in Figure 3, and on prior measure development, we have selected a set of crew-coordination measures for the analysis of two-person helicopter flightcrew coordination and teamwork. This subsection discusses the measures that were selected, and the basis for those selections. The full set of measures will be used in flightcrew experiments planned for the second year of the project. A number of the measures have already been used in preliminary analysis and testing over the past year.

In order to test our theories of crew coordination, we will require measures of workload, crew processes (both taskwork and teamwork), and crew performance. Workload measures are relatively straightforward. Either the SWAT or the TLX may be used—the TLX is recommended for ease of use. The performance measures will be specific to the tasks simulated in the experiment. Subjective evaluations by domain experts using ATM-based measures are recommended, as well as mission-specific measures to be obtained from simulator data. The selection of crew process measures is more problematical, especially for measures of teamwork. Figure 4 shows that a wide variety of approaches have been used to measure teamwork processes. These approaches focus on different aspects

of teamwork, measure different levels of detail, and use different data-collection techniques. We recommend use of the CMAQ to collect self-report data on teamwork attitudes—it is the only validated instrument of its type. Recommendations for additional teamwork measures based on observer rating and frequency-count data are discussed in more detail below.

Teamwork Measures Based on Rating and Frequency-Count Data

Two very different approaches to the measurement of teamwork processes are seen in prior research. The first is the use of ratings of crew behavior and teamwork processes by domain experts along a number of different dimensions that are believed to be important for effective crew performance. This approach includes the CTBF, the ATOM, and the ACE. An alternative approach analyzes crew communications and crew actions in more detail and relies on event-frequency counts rather than on ratings by domain experts. The DDD communication and coordination measures and Orasanu's communication-content measures fall into this latter category. We recommend the use of both types of teamwork measures for crew-coordination studies.

Ratings of Teamwork Process. As mentioned in the literature review, several research groups have attempted to classify teamwork dimensions in different categories, and have suggested appropriate corresponding measures. A large degree of repetition and overlap exists among the several categorizations, and the degree of detail varies widely. We focused on two measurement tools that have been empirically validated and have shown a reasonable level of face validity: the CTBF/ATOM classification of teamwork attributes (Dwyer, 1992; Oser et al., 1989), and the ACE classification of crew-coordination basic qualities (Simon, 1991).

Teamwork Attributes (CTBF/ATOM). The CTBF-based ATOM uses a set of five-point anchored scales spanning seven dimensions of teamwork. These Teamwork Attributes (TAs) have emerged from a refinement of earlier versions of the CTBF classification and include the following (Dwyer, 1992):

- *TA1. Team Orientation* refers to the attitudes that team members have toward one another and the team task. It includes the level of group cohesiveness and self-awareness of team membership. Positive team orientation manifests itself in a situation where one or more team members place the goals and interests of the team ahead of personal goals and interests. It may also be evident through the display of trust, team pride, and esprit de corps.

- *TA2. Team Leadership* involves providing direction, structure, and support for the other team members. It does not necessarily refer to a single individual's formal authority. Positive team leadership manifests itself when

a member works with other members to develop them to the full extent of their capabilities.

- *TA3. Communication* includes both procedural and interactive communication. Procedural communication involves a team member's providing information to other members in the prescribed manner using proper terminology and procedure. Interactive communication involves the exchange of information between two or more team members in order to clarify or acknowledge the receipt of information. Positive communication occurs when team members pass on all important information, clarify intentions and planned procedures, acknowledge and repeat messages to ensure correctness, and ensure that messages are received as intended.

- *TA4. Monitoring* refers to observing the activities and performance of other team members. It implies that team members are individually competent and that they may subsequently provide feedback and backup behavior. Positive monitoring occurs when team members observe the performance of other team members to ensure the efficiency of the team, recognize when other team members perform correctly, and consistently keep track of other team members' performance.

- *TA5. Feedback* involves the giving, seeking, and receiving of information about other team members. Giving feedback refers to providing information regarding other members' performance. Seeking feedback refers to requesting input or guidance regarding performance. Receiving feedback refers to accepting positive and negative information regarding performance. Positive feedback behavior occurs when team members go over procedures with other team members by identifying mistakes and potential corrective action, ask for input regarding their own mistakes, and incorporate suggestions into their behavior.

- *TA6. Backup Behavior* involves assisting the performance of other team members. This implies that members have an understanding of other members' tasks (mutual mental models). It also implies that team members are willing and able to provide assistance to each other. Positive backup behavior occurs when—if a team member is having difficulty, makes a mistake, is becoming overloaded, or is unable to perform duties—other members step in to assist and ensure that the activity is completed correctly.

- *TA7. Coordination* refers to team members' executing their activities in a timely, synchronized, and integrated manner. Positive coordination behavior occurs when team members consistently pass tactical information to team leaders and other team members, thereby enabling them to accomplish tasks, and when team members consistently carry

out their tasks in a timely manner, enabling other team members to carry out team tasks effectively, thereby minimizing the occurrences of team failures.

The seven CTBF/ATOM attributes of teamwork were developed for generic team task environments. It is up to the analysts and experimenters to tailor these dimensions into specific measures and associated scales relevant to the particular task at hand. For example, the ATOM was adapted to measure teamwork in a shipboard Combat Information Center (CIC) within a hierarchical team consisting of a tactical officer and four subordinates.

Team Basic Qualities (ACE). The ACE dimensions or basic qualities (BQs) were developed specifically to evaluate crew coordination and cockpit resource management for flightcrews. The ACE classification includes thirteen dimensions relevant to crew coordination (see previous description), briefly outlined in the following list:

- *BQ1. Establish and maintain flight team leadership and crew climate* includes leadership style, professional respect, resolution of disagreements, and crew member attitude.

- *BQ2. Premission planning and rehearsal accomplished* refers to premission flight planning, premission rehearsal, and in-flight replanning and rehearsal.

- *BQ3. Selection of appropriate decision-making technique* includes separate decision-making strategies to be adopted in conditions of high time-stress, as opposed to low or moderate time-stress.

- *BQ4. Prioritize actions and distribute workload* includes behaviors that support task prioritization in the cockpit and workload distribution among crew members.

- *BQ5. Management of unexpected events* refers to the crew preparation and composure in times of crisis, and appropriate resource management.

- *BQ6. Clarity, timeliness, relevance, completeness and verification of statements and directives* includes adequacy and timeliness, clarity, and acknowledgment procedures.

- *BQ7. Maintenance of mission situational awareness* includes the awareness level of the crew, as well as the awareness of factors that inhibit attention.

- *BQ8. Decisions and actions communicated and acknowledged* refers to the communication of decisions and actions as well as the practice of clarification and acknowledgment among crew members.

- *BQ9. Supporting information and actions sought from crew* includes both the solicitation of crew input and the solicitation of crew assistance.

- *BQ10. Crew member actions mutually cross monitored* includes the crew member's scanning for crew error and the application of the two-challenge rule (back-and-forth acknowledgment).

- *BQ11. Supporting information and actions offered by crew* refers to the anticipation and offering of required information, as well as the anticipation and offering of required assistance among crew members.

- *BQ12. Advocacy and assertiveness* refers to the practice of advocacy: the extent to which crew members advocate a course of action they consider best, even if it may differ from the one being followed or proposed.

- *BQ13. Crew-level after action reviews accomplished* includes the critique and improvement of crew performance.

The ACE decomposes each of the thirteen basic qualities into two to four rating factors, for which an anchored scale is developed. Each scale describes in detail what constitutes superior, acceptable, and poor crew behavior for that basic quality.

Comparison of the CTBF/ATOM and the ACE. We reach two overall conclusions from examination of the CTBF/ATOM and the ACE. First, while the CTBF/ATOM focuses exclusively on teamwork, the ACE includes several taskwork dimensions. These taskwork attributes, such as the maintenance of situational awareness (BQ7), are essential components of crew processes. In fact they may directly result from superior teamwork behaviors, such as the proper exchange of information among the crew members. The following ACE basic qualities can be classified as taskwork-type dimensions: BQ2 (premission planning), BQ3 (decision-making techniques), BQ5 (management of unexpected events), BQ7 (situational awareness), and BQ13 (after-action review). Since our focus is on teamwork, we will limit our comparisons to the remaining teamwork dimensions.

A second conclusion is that there is a significant degree of overlap between the taxonomies used in the CTBF/ATOM and the ACE, as shown in Table 1, but little one-to-one mapping along teamwork dimensions. There is no dimension that appears in one taxonomy but not in the other, and a number of instances in which a single dimension in one taxonomy maps into several dimensions in the other. We conclude that either instrument could be used to measure teamwork process for our purposes without a substantial

Table 1

Teamwork Dimensions in the ATOM/CTBF and the ACE

	ACE BASIC QUALITIES (TEAMWORK DIMENSIONS ONLY)							
CTBF/ATOM TEAMWORK ATTRIBUTES	BQ1 Leadership & Climate	BQ4 Priority & Workload Distribution	BQ6 Statements & Directives	BQ8 Communic & Acknowledg	BQ9 Info & Actions from Crew	BQ10 Cross-Monitoring	BQ11 Info & Actions by Crew	BQ12 Advocacy & Assertiveness
TA1 Team Orientation	X							
TA2 Leadership	X							X
TA3 Communication				X	X		X	
TA4 Monitoring						X		
TA5 Feedback			X			X		
TA6 Backup		X						
TA7 Coordination					X		X	

loss of information.[1] The ACE was developed specifically for flightcrew teamwork measurement, however, so it may have an advantage for studies of flightcrew coordination.

Teamwork Measures Based on Communications Data

Measures of crew coordination should allow for the identification of *communication and coordination patterns* that are associated with *crew errors and error propagation*. For this purpose ratings of crew-coordination "quality" such as those in the ATOM and the ACE, or the presence or absence of behaviors such as those on the CTBF, are not sufficient. We need *quantitative measures* of crew-coordination behavior that are at a sufficient level of detail to identify the *specific aspects* of communication and coordination that are associated with errors. Measures in the spirit of those used in the DDD research and in Orasanu's work are more likely to fulfill the requirement for quantitative measurement.

Communications data may be collected and analyzed at many possible levels of detail. Figure 5 shows a hierarchy of possible levels for communication measurement in a team. At the coarsest level of measurement, requiring the fewest resources, we can

[1] In a future experiment we plan to collect CTBF/ATOM measures and perform a correlational analysis with the ACE basic qualities to confirm this hypothesis.

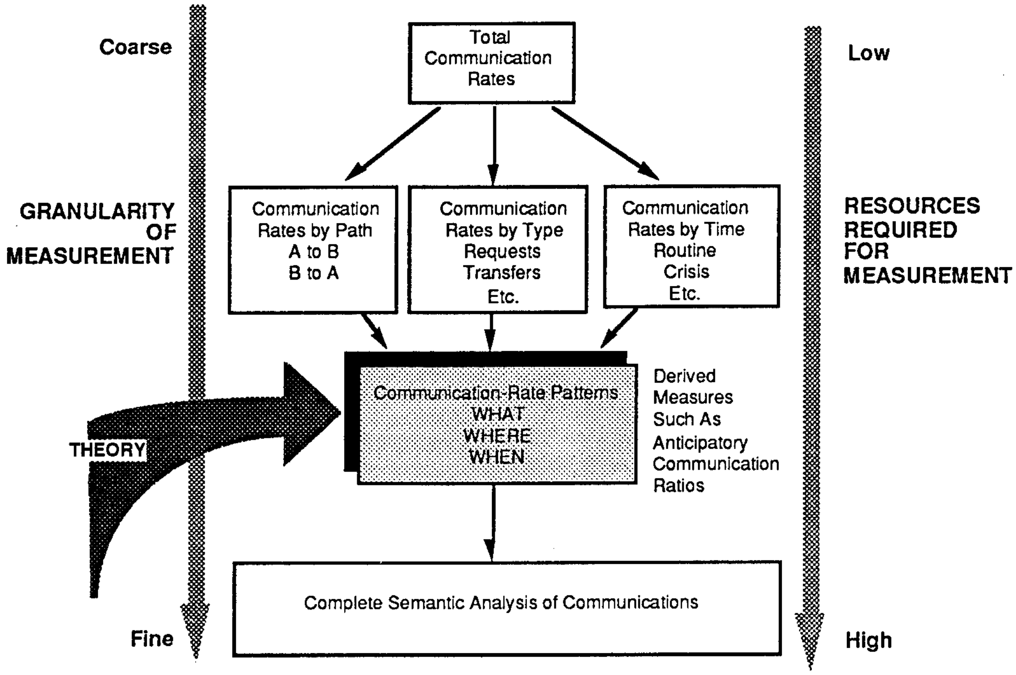

Figure 5. Levels of communication measures.

measure overall communication rates independently of the content of the communication, the individuals involved, or the time at which the communication occurred. If more detail is needed, we can measure communication rates by the type of communication (based on content analysis), or by the path of the communication (the originator and recipient of the communication), or by the time period in which the communication occurred. At the most detailed level, we can perform a complete semantic analysis that considers, for every utterance, the nature of the utterance, the communication path, and the time of occurrence. Such an analysis is typically very resource-intensive and generates an extremely large volume of data.

To balance the need for detail and the economy of resources for analysis, the use of measures of communication-rate *patterns* at an *intermediate* level of granularity is recommended. As shown by the highlighted box in Figure 5, intermediate-level measures focus on the frequency and distribution of communications for *certain selected combinations* of path, content, and time. The patterns on which one should focus are suggested by underlying concepts and theories about the nature of crew coordination and the factors that affect it. For example, the hypothesis that teams practice implicit coordination based on shared mental models suggests that anticipatory behavior is likely to be associated with higher levels of team performance, especially under high time-pressure conditions. A measure of anticipatory behavior, the *anticipation ratio*, may be derived from communications-frequency

measures, based on the ratio of the number of *transfers* of information or resources from team member A to team member B to the number of *requests* of information and resources from B to A.

Summary of Recommended Measures

The measures and measurement tools recommended for analysis of crew coordination in helicopter flightcrews are summarized in Figure 6. We based the selection of the measures and the measurement tools on the information needed to complete the model described in Figure 3. Figure 6 shows the relationship between the components of the model and the sources from which measurements of these components can be derived. We recommend the use of existing measurement tools (e.g., the ATM, ACE, CMAQ, and TLX) and the development of additional tools as necessary. New instruments will definitely be needed for the communication analysis and, depending on the nature of the task, evaluation questionnaires in addition to the ACE may be needed for taskwork/teamwork measurement by domain experts.

<table>
<tr><th rowspan="2">MEASURE / METHOD</th><th rowspan="2">WORKLOAD</th><th colspan="2">CREW PROCESSES</th><th rowspan="2">PERFOR-MANCE (OUTCOME)</th></tr>
<tr><th>TASKWORK</th><th>TEAMWORK</th></tr>
<tr><td>SELF-REPORT (CREWS)</td><td>TLX</td><td></td><td>CMAQ</td><td></td></tr>
<tr><td>AUTOMATED (SIMULATOR)</td><td></td><td>MISSION-SPECIFIC INDICES</td><td></td><td>MISSION-SPECIFIC INDICES</td></tr>
<tr><td>ON-SITE/OFF-SITE OBSERVATIONS (OBSERVERS)</td><td></td><td></td><td>COMMUNI-CATION ANALYSIS</td><td></td></tr>
<tr><td>ON-SITE OBSERVATIONS (DOMAIN EXPERTS)</td><td></td><td colspan="2">ACE
EVALUATION QUESTIONNAIRES</td><td>ATM, SUBJ. EVAL.</td></tr>
</table>

Figure 6. Matrix of recommended methods and measures.

The potential of this measurement approach relies on the theory-predicted links between the different dimensions of crew processes (such as teamwork and coordination), the crews' strategies of adaptation to workload, and the crew performance as evaluated on-line by experts. By correlating the measures resulting from our communication analysis with a combination of measures such as the ATM, the ACE, and the CMAQ, we intend to establish a clearer picture of what constitutes inferior and superior coordination strategies.

A subset of these measurement tools, including a communication-analysis instrument, was used during the first year of the project to analyze the communication and coordination patterns of helicopter flightcrews during simulated missions. A detailed description of this effort is provided in the section that follows.

Detailed Report on Task II-B: Observation and Analysis of Fort Campbell Videotapes

The second major task for the first year of the project was to test—through empirical observation—the measurement approach developed for team coordination, the theory behind the approach, and the measurement instruments based on the theory. We had an opportunity to perform this initial testing through the use of materials from a team-coordination training and assessment project conducted by Dynamics Research Corporation (DRC) under the sponsorship of ARIARDA. We were able to obtain videotapes of helicopter flightcrews during simulated flight before and after they received coordination training at Fort Campbell, Kentucky, and we used these videotapes to test a communication-analysis data-collection instrument. We were also able to obtain a variety of attitude and performance data collected by DRC, which, in combination with measures based on our communication analysis, allowed us to test several aspects of the theoretical framework shown in Figure 3.

This section reports on the results of this observation and analysis of the Fort Campbell videotapes. First, we review the motivation for the analysis and some of the specific hypotheses regarding team coordination and adaptive team behavior that we sought to test. Next, we provide some background on the DRC research effort that produced the videotapes and associated data. We then describe the methodology we used to collect the communications data, including the design of the observation instruments and the rating procedures used. Finally, we present the results of the analysis and discuss its implications for communication and coordination measurement in future research on crew coordination.

Motivation and Key Issues

A majority of aircraft accidents due to human error have been traced to breakdowns of crew coordination and communication in the cockpit during times of crisis. Deficiencies in team interactions, as opposed to poor individual skills, have been the dominant cause of error-making with subsequent catastrophic consequences (Diehl, 1991: Lautman & Gallimore, 1987). To minimize the occurrence of such tragedies, it is essential to understand and measure rigorously the links between particular crew-coordination strategies and superior performance, as well as to identify coordination strategies that are likely to lead to errors. Because most critical errors happen in periods of high workload, it is also important to understand the coping mechanisms

that superior crews use to adapt to increased workload while maintaining their effectiveness (Serfaty, Entin, & Volpe, 1993a). One of our main objectives in this first year of research was to perform initial testing of rigorous methods for measuring crew-coordination patterns and to assess their relationship to crew performance and crew errors. In this subsection we briefly review selected issues pertaining to the relationships between team performance, crew adaptation and coordination strategies, and shared mental models.

Team Performance, Errors, and Adaptation

Key findings from research on team performance can be aggregated into principles with implications for measuring crew coordination (Carley, 1990; Kleinman, Pattipati, Luh, & Serfaty, 1992; Kleinman & Serfaty, 1989; Serfaty, Entin, & Deckert, 1993), and these principles may be used as guidelines to develop measures and propose hypotheses for crew behavior in complex situations. For example, specific team behaviors have been identified as resulting from high-stress or high-workload conditions: as the external load on a team increases, the team members tend to filter out lower-priority tasks, select processed information over raw data, use less explicit coordination (communication), reduce their planning horizons, and ignore tasks with high coordination requirements. As load increases further, teams decompose into a collection of individuals, task load is shared unequally to reduce coordination, team performance degrades significantly, and the rate of errors occurring in the team often results in catastrophic consequences.

A number of issues arise in specifying what is meant by team errors. Salas, Dickinson, Converse, and Tannenbaum (1992) point out that in order to achieve a superior level of performance, efficient and reliable teams must: 1) *coordinate* their resources, information, and actions; 2) *adjust* continuously to the demands of the task environment; and 3) use an organizational *structure* that supports the team process. We suggest that any failure to consistently perform these three activities will result in team errors. Changes in external and internal conditions produce two kinds of errors: *individual errors*, which tend to propagate in the team and affect team performance as a whole; and *team errors*, which occur because of a breakdown or lack of communication in the team.

Although the occurrence of these errors can be significantly reduced by individual training, further improvement can be achieved by: *team structuring* (e.g., organizational structure, information structure, task-responsibility structure) to make team performance more robust to a range of operational conditions; *team adaptation* to changes in the task as well as in the team environment; and *team training*, in which team members are trained specifically in the responsibilities of other team members (cross training) or in communication techniques or other skills over and above training on their individual tasks (Swezey & Salas, 1992).

Observation of teams with records of highly reliable performance suggests that the ability to adapt to a changing environment, with an associated requirement for coordination, may lie at the heart of a team's resilience to errors. La Porte and his colleagues (La Porte & Consolini, 1988; see also Pfeiffer, 1989, and Reason, 1990) have conducted extensive observation and analysis of several types of highly reliable teams operating in real-world environments. Their observations about the common characteristics and procedures of these teams emphasize the importance of flexibility and adaptability in preventing errors in dynamic environments. La Porte identified the following features as characteristic of highly reliable teams:

- The team structure is *adaptive* to changes in the task environment. The reliable team has not one but several organizational structures, and shifts between them as needed. La Porte distinguishes three authority structures: routine, high-tempo, and emergency, each with a different set of characteristic practices, communication pathways, and leadership patterns. The same individuals may play completely different roles under different circumstances.

- The team maintains *open and flexible communication* lines. An adaptable team structure seems to promote the free flow of information from lowest to highest levels as well as the other way around. This is critical in situations where lower levels in a hierarchy may have critical information that is not available to the upper levels. A number of accidents have occurred in which an aircraft's copilot or a ship's crew were aware of a problem but were reluctant to bring it to the attention of the pilot or captain (Green, 1990; Pfeiffer, 1989).

- Team members are extremely *sensitive to other members' workload and performance* in high-tempo situations. La Porte observed that air traffic control teams are very sensitive to the overloading of any team member and, without any overt request for help, will gather around the screen of an overloaded individual to watch for danger points until the overload condition has passed.

Our central hypothesis concerning *team adaptation* evolves from these observations. Superior crews, when faced with a changing—and increasingly demanding—task environment (e.g., crisis vs. routine), will adapt their decision-making strategies, coordination strategies, and even their structure in order to maintain the required level of performance while holding workload at an acceptable threshold (see Figure 3).

<u>Shared Mental Models and Crew Coordination</u>

Findings on adaptive team performance may be explained by a common theoretical premise: effective crews develop a mental model of their common task that enables them to use the team

structure to maintain team coordination and performance under a wide range of conditions. It has been suggested by various authors (Cannon-Bowers, Salas, & Converse, 1990; Kleinman & Serfaty, 1989; MacIntyre, Morgan, Salas, & Glickman, 1988; Orasanu, 1990) that highly effective teams have a *shared mental model* of the situation and the task environment (consistent with the team's understanding of the situation), and a *mutual mental model* of interacting team members' tasks and abilities that generates expectations about how other team members will behave. These mental models are particularly useful in the absence (or scarcity) of timely, error-free, and unambiguous information. We hypothesize that crews that have developed a high level of congruence between their mental models—both situational and mutual—are able to make use of these models to anticipate the way the situation will evolve as well as the needs of the other team members. These crews will perform consistently better under a wide range of flight conditions.

The concept of shared mental models has proved to be a powerful mechanism for understanding the relationship between team-coordination behaviors and team performance. The differences found by Orasanu (1990) in the communication patterns of low-performance and high-performance flight crews suggest that high-performance crews are more successful in developing a shared mental model of their situation. For example, the captains of high-performance crews produced more communications concerning plans and strategies than the captains of the low-performance crews. Under high-workload conditions, first officers in the high-performance crews provided more information to their captains than they did under low-workload conditions, while the pattern is just the opposite for the low-performance crews. Orasanu interprets these findings as indicating that a shared mental model was developed when the captains in the high-performance crews shared their plans, allowing the first officers in those crews to provide information relevant to the plan (i.e., implicitly coordinate) under high-workload conditions.

The coordination mechanisms that support adaptation may be *explicit*, based on specific communications, or *implicit*, based on a shared mental model. Both explicit and implicit coordination will generate observable communication patterns—the presence *and* the absence of communication may be important. For example, communications that provide information to a team member in the absence of requests for that information indicate an implicit coordination mechanism at work. Measures must be sensitive to changes in the team's coordination and communication patterns as they adapt their behavior to the demands of the task and the environment. We expect to see teams shift between explicit coordination (under low-workload conditions) and implicit coordination (under high workload). Even though a team may have a shared mental model that supports implicit coordination, some explicit exchange of information will be required in order to maintain that shared model as the situation changes (Orasanu & Fischer, 1992). Observation has also suggested that members of

highly reliable teams are aware of the workload of other team members and *implicitly* assume some of the tasks of any individual in the team who is overloaded under stressful conditions. This dynamic reallocation of workload should be observable from the team's communication patterns, and our measures should be sensitive to it.

Serfaty, Entin, and Volpe (1993a, 1993b) have proposed a quantitative link between shared mental models and implicit coordination strategies in a team. Based on the premise that a shared mental model would allow team members to implicitly coordinate their actions, and to anticipate the leader's needs for information, they defined an upward-anticipation measure, computed as the ratio of subordinates' information transfers to the leader's requests for information. Results show that the teams tested were *adaptive* to increasing levels of stress. Error rates were higher in the moderate time-pressure condition than in the low time-pressure condition, but were *lowest* in the high time-pressure condition. Teams apparently found a strategy that allowed them to perform accurately under high time pressure: the teams increased their use of implicit coordination based on a shared mental model as time pressure increased. The upward-anticipation ratio measure, an indicator of implicit coordination in hierarchical teams, supports this conclusion. Serfaty et al. found that the anticipation ratio was almost twice as large during high-time-pressure conditions as under low or moderate time pressure, indicating that subordinates were sending information to the team leader without specific requests for that information.

Preliminary results of a similar experiment (Kalisetty, Kleinman, Serfaty, & Entin, 1993) establish a link between *implicit coordination* and performance. A high correlation was found between the upward-anticipation ratios for six teams and their performance levels as measured by the percent of correct responses. Furthermore, as hypothesized, the use of implicit coordination by the highest-performing teams was *adaptive* to time pressure.

These findings on adaptive team performance have led to a theory of team coordination (see the discussion of Figure 3) and to measures based on that theory. One of our primary goals in this initial phase was to test this measurement approach and the theory behind it.

Background of Fort Campbell Observation and Analysis

We were able to test our measurement approach and test portions of our theory of crew coordination through the use of videotapes and data obtained from a research program conducted by Dynamics Research Corporation (DRC) in coordination with ARIARDA (Simon & Grubb, 1992; Wagner, Simon, & Leedom, 1990). This subsection describes this program and the data that we obtained from it.

Background of Program

DRC is conducting research and development of a new training and evaluation program for Army crew coordination. In previous research (Wagner, Simon, & Leedom, 1990), DRC developed a reliable, objective, and quantitative set of metrics for measuring and assessing crew coordination. These measures were developed to assist the Army in evaluating the effectiveness of its Aircrew Coordination Training (ACT) program in improving both crew coordination and crew performance. Using these materials as a basis, DRC developed a methodology and materials for evaluating crew performance.

In a research program conducted at Fort Campbell in 1992, DRC developed an Aircrew Coordination Evaluation Testbed to validate the materials and methodology developed by DRC and to validate the crew-coordination training course developed by the Army. Under its contract with ARI on structuring and training highly reliable teams, ALPHATECH was given access to the materials from the Testbed, including videotapes of crews during simulated helicopter flight sessions, and observational rating and attitude data collected by DRC as part of the Testbed sessions.

Data from the Aircrew Coordination Evaluation Testbed

The Testbed developed by DRC includes the UH-60 Visual Flight Simulator (VFS), mission scenarios, cameras for videotaping of crew performance in the scenarios, and evaluation materials to be collected in pre- and post-training simulations. This subsection summarizes information about the Testbed that is important for understanding the results of this report. A complete description of the Testbed is contained in Simon and Grubb (1992).

The data used in our observation and analysis were collected in August 1992, when 16 flight crews at Fort Campbell participated in the Aircrew Coordination Training course. Prior to the training course, each of the 16 crews completed a pre-training mission in the VFS. After completing the course, each crew performed a post-training mission in the simulator. Each of the pre- and post-training missions was videotaped. Crew-coordination behavior and performance before and after training were evaluated by trained IPs, and mission performance data were recorded in the simulator. The individual attitudes of the crew members were assessed through self-report questionnaires before and after training.

The scenarios used in the pre- and post-training missions were "developed to realistically illustrate the primary missions, conditions, and situations needed to evaluate crew coordination mission performance areas and tasks" (Simon & Grubb, 1992, p. 15). The scenarios included a cross-FLOT (forward line of troops) air assault mission and an external load air movement mission. Each scenario included three crisis situations—a major and a minor aircraft system malfunction and an inadvertent entry into

instrument meteorological conditions (IMC). The simulated mission was performed by a two-person cockpit crew (the subjects) supported by a crew chief located in the rear of the aircraft and by ground personnel (experiment confederates). The mission required about 60 minutes to complete. Actions of the crew, such as getting lost or taking an unusually large amount of time to solve a problem, could lengthen the amount of time needed to complete the scenario.

The primary evaluation measure used to assess crew coordination in the Testbed was the ACE, which "measures an aircrew's ability to integrate a variety of human factors principles into the cockpit milieu" (Simon, 1991, p. 95). The ACE was completed by the IPs responsible for training the crews and for evaluating their performance in the simulator.

DRC developed three types of evaluation measures to assess crew effectiveness. These performance-evaluation measures were provided by the IPs and by the simulator. The first measure was an overall grade, assigned by the IP and based on the IP's overall assessment of the crew's performance on the mission. The grade can have one of four values: U, S-, S, and S+ (coded as 0, 1, 2, and 3, respectively), where U stands for Unsatisfactory and S for Satisfactory. The second performance-evaluation measure was a set of gradeslips based on 18 Aircrew Training Manual (ATM) measures. The ATM measures assess a crew's performance on specified tasks. The gradeslips were filled out by the IPs, who used the same evaluation scale used for the overall grade. The third performance-evaluation measure was a set of mission-performance scores that were designed to link ATM tasks to overall mission performance. These mission-performance measures evaluated six performance areas: terrain flight navigation (based on deviations from the flight plan), threat avoidance/evasion (based on actions during and outcomes of threat encounters), response to aircraft emergencies (based on evaluation of the crew's diagnosis of the emergency, their adherence to emergency procedures, and the outcome of the emergency), response to unexpected events (based on adherence to recovery procedures and on the eventual outcome), instrument recovery (including both approach planning and approach execution), and mission-threatening crew errors (based on an evaluation of the occurrence and adverse effects of crew errors). Between two and five specific performance sub-measures were associated with each of these areas. In addition, the percentage of mission objectives completed during the simulated scenario was computed. The mission-performance scores were based on data from the simulator or were collected by the IPs during the simulated flight (see Simon and Grubb (1993) for a detailed explanation).

DRC used a modified version of the CMAQ to assess crew members' attitudes toward team coordination. The instrument was filled out by each of the crew members both prior to and after they had taken the ACT course. This version of the CMAQ includes 46 items designed to assess crew members' attitudes toward various

aspects of cockpit behavior. The response to each item is based on a 7-point scale.

Analysis of the Testbed Data

DRC has conducted an extensive analysis of the materials and methodology used in the Testbed. Their results and conclusions are reported in Simon and Grubb (1993). In general, DRC concludes that the IPs can reliably perform the ACE and gradeslip evaluations and that these instruments provide valid measures of crew coordination and performance.

With the approval ARI, DRC provided the Testbed data to ALPHATECH to support our development of crew-coordination measurement instruments and our analysis of the mechanisms of crew coordination. A key issue for our research is the effect of the coordination training provided to the crews on coordination and communications patterns, and the links between coordination behaviors and crew performance. A critical assumption in analyzing the communications data before and after training is that the coordination training provided to the crews improved their performance.

We verified this assumption by analyzing differences in pre-training and post-training performance and coordination, as detected by the ACE, the ATM-based measures, and the CMAQ. A summary of these analyses is presented in Appendix B. The analyses indicate significant differences between pre- and post-training on both behavioral and performance measures. The average ACE score (obtained by averaging the eight component scores directly related to crew coordination increased significantly after training, indicating that crew coordination behavior as assessed by the IPs improved after the crews participated in the coordination training. Because we are particularly interested in the nature of the behavioral components that relate to crew coordination, we compared the 13 component scores from the ACE individually, and found that all but the 13th component (after-action reviews) increased significantly from the pre- to post-training administration. The two performance-evaluation measures—the overall grade and the average ATM grade—both showed significantly better performance in the post-training than the pre-training simulations. Furthermore, the attitudes of the crew members, as assessed by the 46-item CMAQ showed a significant change from pre- to post-training.

On the basis of our analyses of the data collected by DRC, we observed that the behavior, performance, and attitudes of the crews changed after they participated in the coordination training course. We expected, then, that our communication-analysis instrument would be sensitive to these changes, and would give us insight into the ways in which communication rates and patterns changed as a result of the training. We also anticipated that the eight components of the ACE that assessed teamwork-related

behavior would be related to measures derived from our communication-analysis instrument.

Objectives of Videotape Observation and Analysis

We had two major objectives in observing the videotapes of the pre- and post-training flightcrews and analyzing the resulting data:

- To test the set of crew coordination and communication measures and the associated measurement instruments that had been developed.

- To familiarize ourselves with coordination patterns in a cockpit environment, and to obtain empirical evidence supporting key hypotheses about the relationships between observable crew-coordination patterns and crew performance.

These goals are not independent—one of the ways in which we assess the value of our crew-coordination measures is by their success in producing valid and meaningful results in support of the theory. The use of videotapes from the Fort Campbell flight simulation experiment provides us with an environment to accomplish both objectives.

Several aspects of the communication-analysis instrument were tested. First, we tested *reliability* across observers—do several observers, viewing the same behaviors, produce similar measurement results? We were also concerned with the *validity* of the communication-analysis instrument and the coordination measures based upon it. Do our measures actually capture the behaviors that we claim they capture? As we used the communication-analysis instrument to record the behavior on the videotapes, we were able to form an impression about the extent to which we were measuring the behaviors that we intended to measure. The second test of the validity of the instruments is based on whether the results obtained are consistent with the theory that guided instrument development. Do the coordination measures show the relationships that we expected with each other, and with outcome measures of crew performance? If so, then we can have some confidence that they captured the intended behaviors. Another test is the *sensitivity* of the coordination measures—do we see the variations that we expected to observe when task demands increase (e.g., routine vs. crisis situations)?

For our second objective, we examined the relationship between crew performance and the strategies used by crews to coordinate their activities. Because the scenarios we observed were not under our control, we could not test our hypotheses directly. Instead, we looked for evidence of certain hypothesized behaviors and relationships. For future experiments, it will be important to create more-controlled conditions that motivate teams to adapt their behavior to the demands of the task.

The analysis of the videotapes from the Fort Campbell experiment is an exploratory study lacking a systematic experimental design. It capitalizes on the availability of empirical data collected by the DRC team (Simon and Grubb, 1993), however, and on their analysis of the ACE, ATM, and CMAQ measures. To complement the DRC study, we take that analysis a step further by:

- Performing a *quantitative* analysis of the communications data that provides a more precise evaluation of the impact of the crew-coordination training. For example, if the subjective evaluation (ACE checklist) found that intra-crew communication tends to increase as a result of coordination training, our analysis will quantify and qualify that increase, i.e., answer the questions: how much, and what.

- Performing a *detailed* analysis of the crew communications components and their effect on implicit and explicit coordination strategies used by the flight crews. For example, our analysis decomposes the communication components into functional elements (e.g., information, action, and planning) and quantifies the contribution of each element to the overall crew communications. In addition, indicators of implicit coordination strategies (e.g., anticipation ratios) are analyzed under both crisis and routine conditions to examine the patterns of crew adaptation to stressful situations.

Our observation and analysis plan has five principal components:

1. Analysis of differences in crew communication/coordination patterns between pre- and post-training phases,

2. Analysis of correlations between the eight ACE Basic Qualities related to teamwork and our indicators of crew communication/coordination,

3. Analysis of relationships between CMAQ attitudinal components and our measures of communication/coordination,

4. Analysis of correlations between our measures of crew communication/coordination and crew performance measures, including:

   - Analysis of correlations between performance and coordination for both pre-training and post-training phases.

   - Analysis of differences between high- and low-performance crews, and

5. Recommendations for refining the battery of team measurement instruments.

This preliminary observation of videotaped flightcrew behavior contributes to three main goals of the project: 1) rigorously and quantitatively *measure communication patterns* in flight crews; 2) *link crew-coordination strategies* to team performance; and 3) *trace the occurrence of team errors* to specific coordination patterns. The results of the observation will support the development of crew-coordination measures that capture critical communication behaviors. Future research will continue to combine analysis of intra-crew communications data with the use of subjective measures of crew performance and crew-process evaluation by domain experts to establish the relationship between crew-coordination strategies and superior crew performance.

Methodology

This subsection describes the methodology used to obtain communications data from observation of the simulated-flight videotapes, including the preparation of the data-collection instruments, the conduct of the observation, the analysis of inter-rater reliability, and the derivation of measures for analysis.

Preparation of the Coding Instruments

Two types of instruments were developed—a team-communication coding instrument, and team-coordination evaluation instruments. Appendix C contain copies of the instruments.

Team-communication coding instrument. The purpose of the communications analysis was to capture the nature and direction of the communication flow between the two members of a helicopter cockpit crew. Our intent was to capture communication-rate *patterns* at an *intermediate* level of granularity, as suggested by our theory of team coordination (see Figure 5). The measurement instrument we developed permits us to record significant aspects of the crew's communication patterns, but it does not require a complete semantic analysis of the cockpit dialog. The instrument provides a *quantitative* record of the type, function, and directionality of communications as well as the condition (routine or crisis) in which the team is operating and the time period in which the communications occur.

Figure 7 shows the recording instrument developed for data collection, with data from a typical 15-minute segment of a flight session marked on the form as an example. The instrument captures four dimensions of communications: type, function, directionality, and condition. The rows delineate the nature of the communications. Guided by theories of crew coordination, we divided the communications into two types—requests and transfers. Requests include information or actions solicited by one member of the crew. Transfers are statements, including observations of events, statements about actions completed or intended, and acknowledgments. Within each type of communication there are

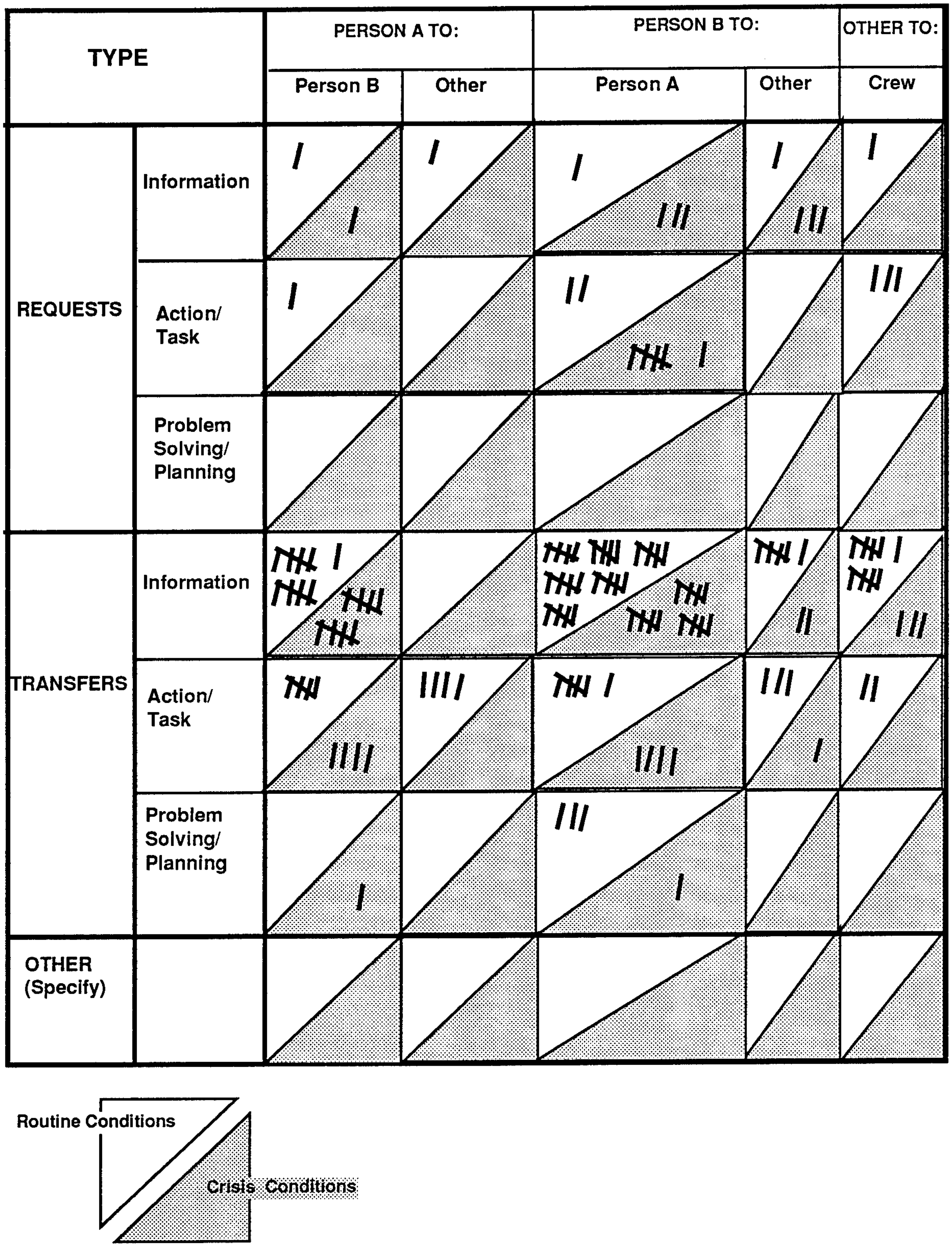

Figure 7. Coding instrument used to capture communications between crew members (data for a 15-minute segment of a flight session shown as an example).

three functional areas: 1) information, 2) action or task, and 3) problem solving/planning. The columns record the direction of communication flow. Because we were primarily interested in the interchange between the cockpit crew members, the recording form groups communications to and from other crew members and ground personnel into an "other" category. The diagonal upper and lower halves of each cell are used to record communications made during routine operating conditions (white) and crisis situations (gray).

Raters using this instrument classified each utterance on the videotape into its appropriate category, and put a tally mark in the appropriate cell on the recording instrument. Thus, the observation and rating process produced a set of numeric scores representing the frequency of occurrence of each category of communication.

Team-coordination evaluation instruments. We also devised two instruments that allowed raters to evaluate team coordination. One instrument, the Teamwork Dimension Rating Form, was designed to assess six aspects of teamwork: team orientation, communication behavior, monitoring behavior, feedback behavior, backup behavior, and coordination behavior. It was patterned after the CTBF/ATOM, and, like the ATOM, uses a behaviorally-anchored rating scale. The second instrument, the Rater's Subjective Evaluation Form, was designed to evaluate each of the two cockpit crew members in terms of their ability to anticipate and understand each other's needs and to rate the degree to which the crew adjusted to a varying workload, the degree to which they provided critical information to each other, the extent to which they explained their actions and plans, and the way in which they handled conflict. This instrument also uses a behaviorally anchored rating scale.

Coding Methodology

The first step in coding the communications data was to establish and calibrate a common set of coding procedures for the two raters. We then coded communications by category directly from the videotapes.

Baseline training and calibration. Two raters who were experienced in protocol analysis, but were not domain experts in helicopter flight or operations, performed the cockpit communication coding for the Fort Campbell tapes. In order to train themselves and establish an agreed-upon set of coding conventions, the raters randomly selected one team's pre- and post-training tapes to use for training and calibration. Before they began to code the tapes, the raters listened to a portion of one of the training tapes and agreed upon an initial set of coding conventions. This allowed the raters to clarify the meaning and intention of the coding categories and to establish norms for coding the types of utterances typically used in the task being studied. After establishing the initial conventions, the two

raters viewed several 15-minute segments of the tape and further clarified and refined their coding conventions.

With a baseline set of standards in place, each of the raters coded a 15-minute segment of the tape. They compared their results, looking at both the number of communications coded for each individual (the column sums on the coding sheet) and the categories into which the communications were coded. The raters then viewed the tape together, discussed their differences, and further refined the set of coding conventions to be used. This process was repeated until the raters had achieved at least an 80 percent agreement for two consecutive 15-minute segments, which occurred after three segments. The team used for rater calibration was excluded from the remainder of the analysis, reducing the total number of teams in the analysis from 16 to 15.

Coding the Tapes. After they had completed their training procedure, the two raters coded the 30 Fort Campbell tapes that were used for analysis. Each rater coded half of the tapes. In addition, in order to assess inter-rater reliability, five of the 30 tapes were coded by both raters. To eliminate any biases that might occur from knowing whether a tape was a pre- or post-training tape, the 30 tapes were randomly assigned numbers from one to 30, and the correspondence between these numbers and the team/session codes was hidden from the two raters. The raters worked independently during this phase.

For identification purposes, the crew member who started out flying the aircraft was coded as "person A" and the crew member who started out as the navigator was coded as "person B." In some crews, these individuals switched roles during the course of the mission, but that switch in roles was not captured in the coding forms.[2] To permit comparisons of the communication patterns at different times during a simulated flight, the coding was segmented into time intervals of 15 minutes, with a separate coding form used for each time segment. In most cases the missions were completed within 75 minutes, so there were five coding sheets per tape. In the small number of cases in which the simulation took more than 75 minutes, the fifth time segment included the remainder of the tape. The total simulation time was recorded on the last sheet.

As explained above, the coding was done directly from the videotapes. At the beginning of each tape, before they began coding, the raters listened to a small segment of the tape in order to learn to distinguish the two crewmembers' voices. Once they began coding the tape, it was possible for the raters to code the communications in real time in some segments of the tapes. In other segments—when the audio was muffled, there was a high density of communications, or the semantic context was ambiguous—

[2] Possible modifications to the technique to capture changes in roles are discussed at the end of this section.

the raters needed to listen to a small segment of the tape three to four times in order to insure that their coding was accurate. Overall, in performing the coding of the Fort Campbell tapes, the raters were able to achieve a 2.5-to-1 ratio of coding-time to real-time.

Table 2 shows example utterances for each of the coding categories. The table includes examples of utterances classified as requests for planning/problem solving, but there were very few utterances in the tapes that could be classified into this category. As a result it was eliminated from subsequent analysis.

After they had finished the communication coding of a tape, the raters filled out the two team-evaluation instruments. These two instruments required approximately five minutes to complete.

Assessment of Inter-Rater Reliability

We assessed the reliability of our coding process using coefficient alpha, a standard measure of inter-rater reliability (Cronbach, 1970; Nunnally, 1967). The value of coefficient alpha can vary from zero, which indicates no agreement, to one, which indicates perfect concordance between two raters. We computed coefficient alpha between the two raters for each coding category for each time period. Because our focus is on communications in the cockpit, we used only the within-crew communications. By analyzing specific segments of data (i.e., each individual coding category) rather than summing across categories or time, we took a stringent approach to our assessment of inter-rater reliability.

Our analysis generated a set of 50 coefficients based on the five tapes coded by both raters. The mean coefficient alpha was .82 (standard deviation = .24). Only 10 percent of the coefficients were less than .50, and more than half were over .90. Table 3 shows the coefficient alpha values averaged across the five time periods. The values are very high, with only two of the 10 average coefficients falling below .80. The inter-rater reliability analysis confirmed that the two raters were able to categorize utterances into distinct categories and that they were consistent with each other in classifying the utterances.

We also performed a reliability analysis for the team-coordination evaluation instruments that were completed by both raters. The resulting coefficient alpha values were low, indicating very low agreement between the two raters. We noted that there was very little variability across the teams in the values assigned by the raters. Both raters used a narrow segment of the rating scale and reported difficulty in assessing the teams on the dimensions used in this instrument. We conclude from this analysis that raters who are not domain experts cannot reliably evaluate the *quality* of the teamwork performed by a cockpit crew. Given the low degree of inter-rater reliability and the lack of variability in the ratings, we did not use these evaluations in our subsequent analysis.

Table 2

Examples of Utterances of Each Type and Function

Type	Function	Example
Requests	Information	Do you want me to take it in? What's our hard time? Do you see the river? Did you start a fuel check? [Should we] land to the right of the trees?
	Action/ Task	If you see a building let me know. Keep the high ground on our left. (I'll shoot it.) You just talk to me. Turn to the right. Let's go back to the bridge.
	Problem Solving/ Planning	Let's talk about this a little bit (problem has occurred) Do you think I should call and let them know we will pick up the load? It might help if we found out how far away we are Let me get my headings here.
Transfers	Information	We are clear on the right. I have a house off to the right. Visibility's getting pretty poor. Yes [or no] (said in response to a yes/no question) I think we've come too far.
	Action/ Task	I'm going to cross-feed on number 2. I am going to bump up the air speed. I am going to get back on the river here. Roger. (in response to a request for action)
	Problem Solving/ Planning	We'll need to get fuel. The inbound course will be 289. We'll start the fuel check when we get to .... The check point we're going to be looking for is a road.

Table 3

Average Values for Coefficient Alpha by Category and Direction of Crew Communications

	Direction	
Category	A to B	B to A
Requests - Information	.80	.50
Requests - Action/Task	.74	.83
Transfers - Information	.93	.89
Transfers - Action/Task	.85	.90
Transfers - Planning/ Problem Solving	.88	.95

Preparation of the Data for Analysis

The coding instrument shown in Figure 7 assessed crew communications in terms of four dimensions. (We included a fifth dimension, time period, by using a different recording sheet for each 15-minute time period during the scenario. This dimension was not used in our subsequent analysis.) These four dimensions, and the coding categories for each dimension, are shown in Table 4. To analyze the data, the tallies for each category were summed over the five recording periods. This resulted in a set of 60 frequency counts (two types × three functions × five paths × two conditions). The set of scores for each tape was reduced from 60 to 50 because the raters found very few instances of requests for problem solving/planning. Because we were interested primarily in cockpit communications, we focused the data analysis on communications between the cockpit crew members and eliminated the crew-to-other and other-to-crew communications. Furthermore, because the crew members switched roles as pilot and co-pilot during the scenario, we could not tie communications to role, and therefore summed the A-to-B and B-to-A communications for each category. This resulted in a set of 10 frequency counts that served as the primary set of data, as shown in Table 5.

Based on the 10 basic data categories shown in Table 5, we computed the 17 communication measures shown in Table 6. These measures form the core of our analysis. Based on the hypotheses discussed at the beginning of this section, we were interested in two major aspects of communication. First, we were interested in *anticipatory behavior*, especially in crisis conditions. Anticipatory behavior is measured by the *anticipation ratio*, which divides the number of transfers in a particular category by the number of requests in that category. Our theory suggests that

Table 4

Dimensions of Cockpit Communications

Dimension	Categories
Function	Information Action/Task Planning/Problem Solving
Type	Transfers Requests
Path	A to B A to Other B to A B to Other Other to Crew
Condition	Routine Crisis

Table 5

Ten Basic Data Categories

Requests - Information: Routine Conditions	Requests - Information: Crisis Conditions
Requests - Actions/Tasks: Routine Conditions	Requests - Actions/Tasks: Crisis Conditions
Transfers - Information: Routine Conditions	Transfers - Information: Crisis Conditions
Transfers - Actions/Tasks: Routine Conditions	Transfers - Actions/Tasks: Crisis Conditions
Transfers - Planning/Problem Solving: Routine Conditions	Transfers - Planning/Problem Solving: Crisis Conditions

Table 6

Measures Used for Analysis of Crew Coordination

Communication Measures	Definition
ANTICIPATION RATIOS	
Total Anticipation Ratio	Sum of all transfers divided by sum of all requests
Anticipation Ratio for Information Transmissions	Sum of information transfers divided by sum of information requests
Anticipation Ratio for Actions and Tasks	Sum of action/task transfers divided by sum of action/task requests
Anticipation Ratio in Routine Conditions	Sum of all transfers during routine conditions divided by sum of all requests during routine conditions
Anticipation Ratio in Crisis Conditions	Sum of all transfers during crisis mode divided by sum of all requests during crisis conditions
COMMUNICATION RATES	
Overall Communication Rate—Total (utterances per minute)	Sum across function, type, path, and condition divided by time
Proportion of Communication Rate for Information Transmissions	Communication rate for information transmissions divided by Overall Communication Rate
Proportion of Communication Rate for Actions and Tasks	Communication rate for action/task transmissions divided by Overall Communication Rate
Proportion of Communication Rate for Planning and Problem Solving	Communication rate for planning/ problem-solving transmissions divided by Overall Communication Rate
Communication Rate in Routine Conditions (utterances per minute)	Sum across function, type, and path in routine conditions divided by time in routine conditions

Communication Rate in Crisis Conditions (utterances per minute)	Sum across function, type, and path in crisis conditions divided by time in crisis conditions
Proportion of Communication Rate for Information Transmissions in Routine Conditions	Communication rate of information transmissions during routine conditions divided by Communication Rate in Routine Conditions
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	Communication rate of action/task transmissions during routine conditions divided by Communication Rate in Routine Conditions
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	Communication rate of planning/ problem-solving transmissions during routine conditions divided by Communication Rate in Routine Conditions
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	Communication rate of information transmissions during crisis conditions divided by Communication Rate in Crisis Conditions
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	Communication rate of action/task transmissions during crisis conditions divided by Communication Rate in Crisis Conditions
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	Communication rate of planning/ problem-solving transmissions during crisis conditions divided by Communication Rate in Crisis Conditions

higher-performance teams show more anticipatory behavior, especially under stress. We computed an overall anticipation ratio, anticipation ratios for information transmissions[3] and for actions and tasks, and anticipation ratios for routine and crisis conditions.

We were also interested in the crew *communication rate,* and the *proportion of communications by function*, categorized as information transmission, actions and tasks, and planning and problem solving. These measures are essential for understanding not only how much but also what kind of communication occurred.

[3] In the definition provided in the table we use the term "transmissions" to include both requests and transfers.

Our hypothesis was that higher-performance teams might show relatively more planning-related communications, especially under routine conditions. We computed the overall communication rate by dividing the number of communications by the number of minutes of simulation time. The communication rate thus represents the volume of communications per minute. We then computed the proportion of this overall communication rate related to each of the three functions. Because we expected that communication patterns might be different in crisis versus routine conditions, we also computed the communication rates in both conditions, and the proportion of communications related to each of the three functions in both conditions. These 17 measures are used in our analyses of team coordination presented in the following subsection.

Analysis of Results

This subsection presents the results of our analysis of communication and coordination patterns for the Fort Campbell helicopter flightcrews. We begin by examining the sensitivity of the communications measures (defined as described in Table 6) to the effects of the crew-coordination training, comparing the values for each measure before and after training. We then examine the correlation of the communications measures with the crews' scores on other teamwork-measurement instruments—the ACE and the CMAQ. We then examine the correlations between crew-communications measures and crew performance as measured by the overall grade given to the crew and their scores on the ATM measures. This analysis takes two forms—a correlation of overall grade, average ATM scores, and mission-performance scores with the communications measures; and a contrast of communications measures for the highest-performance versus the lowest-performance teams, defined based on the overall grade and the ATM scores.

Communication and Coordination Patterns Before and After Training

We expected that the coordination training provided to the helicopter crews would change their communication and coordination patterns, and we also expected that our measures would be sensitive to these changes. Table 7 shows the means for each of our communication measures for the 15 teams before they were trained, and the means for the same 15 teams after training, with a paired-sample t-test for the significance of the differences between the means.[4] A number of our measures were significantly different after training, in directions predicted by our hypotheses.

[4] In Table 7 and the tables that follow, significance levels based on one-tailed tests are indicated as follows: ‡ = $p \leq .10$, * = $p \leq .05$, ** = $p \leq .01$. We used one-tailed tests for two reasons. First, in most instances we had strong prior hypotheses about the direction of the effects. Second, this was an exploratory study, and we wanted to err on the side of including rather than excluding variables of interest.

Table 7

Communication Measures Pre- and Post-Training

Communication Measures	Pre-training	Post-training	t-test
ANTICIPATION RATIOS	Mean	Mean	
Total Anticipation Ratio	2.47	2.83	2.27*
Anticipation Ratio for Information Transmissions	4.60	4.48	-0.40
Anticipation Ratio for Actions and Tasks	0.97	1.27	2.19*
Anticipation Ratio in Routine Conditions	2.39	2.71	1.99*
Anticipation Ratio in Crisis Conditions	3.10	3.48	1.08
COMMUNICATION RATES			
Communication Rate—Total	10.03	11.91	3.40**
Proportion of Communication Rate for Information Transmissions	0.57	0.54	-1.49‡
Proportion of Communication Rate for Actions and Tasks	0.36	0.36	.17
Proportion of Communication Rate for Planning and Problem Solving	0.07	0.10	2.32*
Communication Rate in Routine Conditions	10.24	11.96	2.70**
Communication Rate in Crisis Conditions	9.35	11.83	3.35**
Proportion of Communication Rate for Information Transmissions in Routine Conditions	0.57	0.53	-1.86**
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	0.36	0.37	.65
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	0.07	0.10	2.55*
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	0.58	0.59	.34
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	0.35	0.32	-2.11**
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	0.07	0.09	1.16

The overall anticipation ratio (the number of transfers divided by the number of requests) was significantly higher for crews after training, consistent with the hypothesis that *implicit coordination based on a shared mental model produces anticipatory behavior in more-effective crews*. The increase in the anticipation ratio applied only for communications involving actions and tasks, however, not for purely informational requests and transfers. The coordination training provided to the crews apparently increased the extent to which crew members volunteered information about the actions and tasks they were completing without being asked. Anticipation ratios were higher (more anticipation) after training in both crisis and routine conditions, but the difference was significant only for routine conditions. In both pre- and post-training, the anticipation ratios were higher in crisis conditions than in routine conditions, consistent with our hypothesis that crews rely more heavily on anticipatory behavior when they are under pressure.[5] We did not find an interaction between anticipation ratios in crisis and routine conditions and the coordination training, however. That is, we did not find that training significantly *increased the difference* in the anticipation ratio between routine and crisis situations.

The total communication rate was significantly higher (by almost 20 percent) after the training (10.03 communications per minute before training and 11.91 after training), and the *nature* of the communications changed as well. Table 7 shows the proportion of the overall communication rate that was made up of information transmission, discussion of actions and tasks, and planning and problem solving. The proportion of communications related to planning and problem solving was significantly higher after training, and the proportion related to information exchange significantly lower. Crews communicated at a higher rate after training, and proportionately more of that communication involved planning and problem solving and less involved the exchange of information. This is consistent with Orasanu's (1990) finding that high-performance teams produced more communications involving plans and strategies than low-performance teams.

The post-training communication rate was significantly higher in both routine and crisis conditions, but the increase was larger in crisis conditions. After training, the crew communicated at almost the same rate during the crisis and routine conditions, whereas before training they communicated somewhat more during routine conditions. The nature of the pre- and post-training communication changes was different in crisis and routine conditions. In routine conditions, training increased the proportion of communication involving planning and problem solving, and decreased the proportion of communication involving

[5] The t-test for the difference in the anticipation ratio in routine and crisis conditions was 2.52 ($p \leq .01$) pre-training and 5.03 ($p \leq .001$) post-training.

information transmission. In crisis conditions, training decreased the proportion of communication involving actions and tasks. A comparison of the proportion of communications in each of the three functional categories in routine and crisis conditions shows that, after training, there was a higher proportion of communications involving information exchange during crisis than during routine conditions, and a lower proportion of communications involving actions and tasks.[6] This is consistent with our hypothesis that implicit coordination allows the higher-performance team to anticipate actions and tasks under high-pressure conditions, and with the finding that the anticipation ratio for actions and tasks was higher after training.

Overall, we find that the coordination training provided to the helicopter crews resulted in more anticipatory behavior, especially anticipatory behavior involving communications about actions and tasks. Training not only increased the volume of communications but also changed the communication pattern, leading to relatively less exchange of information and more communication about planning and problem solving under routine conditions, and relatively less communication about actions and tasks during crisis conditions.

Correlation of Communication Measures with ACE Basic Qualities

The ACE instrument measures a number of dimensions or basic qualities of a team as rated by observers who are subject-matter experts. One of the goals of our analysis was to determine whether, and to what extent, our communication measures were correlated with the expert observers' ratings of the ACE dimensions. Eight of the ACE dimensions deal primarily with teamwork (see the discussion in the previous section). Tables 8 and 9 show the correlations of these eight ACE dimensions with our communication measures both before (Table 8) and after (Table 9) training. Correlations with a p-value of .10 or less are indicated in the tables, and the last column and last row of each table provide a count of the number of correlations in that column or row that are significant at the 0.10 level or less.

The most striking overall pattern in Tables 8 and 9 is that our communication measures and the ACE basic qualities were much more highly correlated after training than before training. In the pre-training data, there were only 15 significant correlations out of a total of 136. This is little more than one would expect to find by chance at a significance level of 0.10. The most consistent pattern in the pre-training data is the correlation between BQ12, which rated advocacy and assertiveness, and the type

[6] The t-tests for differences between the proportion of communications in each category during routine and crisis conditions after training are $t=-3.22$ ($p \leq .01$) for information exchange and $t=3.13$ ($P \leq .01$) for actions and tasks. There were no significant differences in routine and crisis conditions before training.

Table 8

Correlation of Communications Measures with ACE Basic Qualities Before Training

Communication Measures	ACE Basic Qualities								
	BQ 1 Leadership & Climate	BQ 4 Priority & Workload Distribution	BQ 6 Statements & Directives	BQ 8 Communications & Acknowledgments	BQ 9 Info & Actions from Crew	BQ 10 Cross-Monitoring	BQ 11 Info & Actions by Crew	BQ 12 Advocacy & Assertiveness	Number Significant Correlations
ANTICIPATION RATIOS									
Total Anticipation Ratio	-.07	.26	-.03	.03	-.10	.15	-.12	.08	
Anticipation Ratio for Information Transmissions	-.24	-.15	-.16	-.12	-.20	.02	-.42‡	.27	1
Anticipation Ratio for Actions and Tasks	.02	.27	-.07	-.10	-.20	.12	-.03	.24	
Anticipation Ratio in Routine Conditions	-.05	.23	-.07	-.04	-.15	.05	-.20	.04	
Anticipation Ratio in Crisis Conditions	-.01	.27	.14	.22	-.01	.42‡	.17	.25	1
COMMUNICATION RATES									
Communication Rate-Total	.13	-.02	-.25	-.20	.08	-.23	.03	-.01	
Proportion of Communication Rate for Information Transmissions	-.10	-.03	-.06	.24	.12	-.16	-.33	-.56*	1
Proportion of Communication Rate for Actions and Tasks	-.05	-.15	-.35‡	-.40‡	-.26	-.01	.03	.25	2
Proportion of Communication Rate for Planning and Problem Solving	.09	.23	.03	.16	.15	.23	.43*	.48*	2
Communication Rate in Routine Conditions	.14	-.07	-.29	-.23	.09	-.26	.06	.04	
Communication Rate in Crisis Conditions	.09	.32	.17	.05	-.02	.06	-.10	-.27	
Proportion of Communication Rate for Information Transmissions in Routine Conditions	-.01	-.01	.29	.22	.11	-.20	-.33	-.55*	1
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	-.03	-.22	-.35‡	-.42‡	-.24	.00	.06	.29	2
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	.05	.30	.04	.24	.15	.29	.40‡	.42‡	2
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	-.52*	-.08	.21	.21	.11	.16	-.19	-.33	1
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	.39‡	.26	-.22	-.05	-.21	-.03	-.08	-.01	1
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	.22	-.19	-.02	-.22	.10	-.17	.34	.43*	1
Number of Significant Correlations	2		2	2		1	3	5	15

Table 9

Correlation of Communications Measures with ACE Basic Qualities After Training

Communication Measures	ACE Basic Qualities								
	BQ 1 Leadership & Climate	BQ 4 Priority & Workload Distribution	BQ 6 Statements & Directives	BQ 8 Communications & Acknowledgments	BQ 9 Info & Actions from Crew	BQ 10 Cross-Monitoring	BQ 11 Info & Actions by Crew	BQ 12 Advocacy & Assertiveness	Number of Significant Correlations
ANTICIPATION RATIOS									
Total Anticipation Ratio	.50*	.19	.62**	.22	.34	.19	.34	.54*	3
Anticipation Ratio for Information Transmissions	.56*	.20	.37‡	.15	.22	.10	.22	.48*	3
Anticipation Ratio for Actions and Tasks	.23	-.07	.31	.04	.06	-.04	.05	.40‡	1
Anticipation Ratio in Routine Conditions	.42‡	.21	.57**	.19	.28	.16	.28	.58*	3
Anticipation Ratio in Crisis Conditions	.59**	.02	.52*	-.19	.43*	.16	.43*	.11	4
COMMUNICATION RATES									
Communication Rate-Total	.00	.34‡	.22	.44*	-.00	.16	-.00	.26	2
Proportion of Communication Rate for Information Transmissions	-.04	-.26	.04	-.13	-.03	-.29	-.03	-.11	
Proportion of Communication Rate for Actions and Tasks	-.05	.10	-.19	-.08	-.29	-.11	-.30	.03	
Proportion of Communication Rate for Planning and Problem Solving	.12	.49*	.36‡	.28	.39‡	.52*	.39‡	.11	5
Communication Rate in Routine Conditions	.00	.31	.21	.42‡	-.03	.10	-.03	.24	2
Communication Rate in Crisis Conditions	-.04	.47*	.21	.49*	.15	.39‡	.15	.34	3
Proportion of Communication Rate for Information Transmissions in Routine Conditions	-.04	-.17	.12	-.11	-.01	-.26	-.01	-.04	
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	.07	-.18	-.42*	-.10	-.30	-.10	-.29	-.03	1
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	.14	.48*	.36*	.29	.39‡	.50*	.39‡	.10	5
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	-.11	-.43‡	-.25	-.13	-.12	-.30	-.12	-.21	1
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	.10	.14	.02	-.02	-.15	-.05	-.15	.17	
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	.05	.48*	.34‡	.22	.36‡	.52*	.36‡	.12	5
Number of Significant Correlations	4	7	8	3	4	4	4	4	38

of communications observed. BQ12 was negatively correlated with the proportion of communications that dealt with information exchange, and positively correlated with the proportion of communications that dealt with planning and problem solving in both routine and crisis conditions.

In the post-training data, there were 38 significant correlations between the ACE dimensions and our communications measures. All but one of the eight ACE BQs was correlated with four or more communications measures. The BQ dimension most frequently correlated with communication measures was BQ6, which rated the clarity, timeliness and completeness of statements and directives. This measure was significantly correlated with the overall anticipation ratio, the anticipation ratio for information transmission, and the anticipation ratio in both routine and crisis situations. BQ6 was also positively correlated with the proportion of communications that dealt with planning and problem solving in both routine and crisis conditions, and negatively correlated with the proportion of communications that dealt with actions and tasks in routine conditions.

BQ4, which rates the crew's prioritizing of actions and distribution of workload, was also correlated with a number of communication measures, including a positive correlation with the total communication rate, the communication rate in both crisis and routine conditions, and with the proportion of communications that dealt with planning and problem solving in both crisis and routine situations. It was negatively correlated with the proportion of communications that dealt with information transfer during crisis conditions.

Other BQ dimensions were correlated with fewer communication measures. BQ1, which rates leadership and climate, and BQ12, which rates advocacy and assertiveness, were positively correlated with four of the five anticipation ratios, but not with the communication-rate or communication-type measures. BQ9, which rates the information and actions sought from the crew, and BQ11, which rates the information and action offered by the crew, were positively correlated with the anticipation ratio only in crisis conditions. BQ9 and BQ11 were also positively correlated with the proportion of communication that dealt with planning and problem solving in both routine and crisis conditions. BQ10, which rates crew cross monitoring, was positively correlated with the proportion of communication that dealt with planning and problem solving in both routine and crisis conditions, and with the communication rate during crisis conditions. BQ8, which measures communications specifically, was positively correlated with communication rate measures in both crisis and routine conditions as expected, but not with anticipation ratio or communication function.

Overall, we see that for the post-training data there is a considerable amount of correlation but no one-to-one mapping between specific communications measures and individual ACE BQ

dimensions. Crews that devoted proportionately more of their communications to planning and problem solving were rated more highly on the ACE BQ4, BQ6, BQ9, BQ10, and BQ11 dimensions. Crews that showed higher anticipation ratios were rated more highly on the BQ1, BQ6, BQ9, BQ11, and BQ12 dimensions. Crews with a higher communication rate, regardless of type, were rated more highly on the BQ8 dimension.

The much stronger patterns of correlation between the ACE dimensions and the communication measures after training appears to be caused by differences in the pre-training and post-training variability of both the ACE dimensions and the communication measures. Overall, the crews appeared more similar on both the ACE and the communication measures before they were trained than afterwards. Before training, the mean score for the ACE BQ dimensions was 3.15 with a standard deviation of .55. After training, the mean was 4.33, with a standard deviation of .90. The standard deviation for seven of the eight ACE BQ dimensions shown in Tables 8 and 9 increased from pre- to post-training—the average increase was 29 percent. The variability of the communication measures increased also, but not as much as for the ACE dimensions. For the 17 measures shown in Tables 8 and 9, six had lower standard deviations after training and 11 had higher standard deviations—the average change was an increase of 10 percent. The larger variability in both the ACE dimensions and the communication measures after training resulted in a broader range of scores on both types of measures, apparently producing a stronger pattern of correlations after training.

Correlation of Communications Measures with CMAQ Responses

The CMAQ measures the crew's attitudes regarding teamwork and coordination. The version of the CMAQ administered to the helicopter crews pre- and post-training contains 46 items, scored from 1 (strongly disagree) to 7 (strongly agree). A number of the CMAQ attitude questions are directly related to our hypotheses about the factors that contribute to effective crew performance. Table 10 lists the CMAQ items that we believe, based on visual inspection, to be related to two of our major hypotheses: the hypothesis that highly reliable teams maintain open communications and a flexible division of responsibilities based on a *shared* mental model of the situation; and the hypothesis that highly reliable crews have *mutual* mental models based on sensitivity to the performance and workload of other crew members, and that these models allow them to anticipate the needs of the other crew members. We found 14 CMAQ items that seem directly related, based on their face validity, to the first hypothesis, and combined them into a single index by taking an average. Note that seven of the items are positively related to the hypothesis, i.e., a strong agreement with the item indicates agreement with the hypothesis, while seven are negatively related. For the negative items, we reversed the direction of the scores before combining them with the positive items. We found six CMAQ items that seem, based on face validity, positively related to the second hypothesis, which

Table 10

Derivation of Two Indices from CMAQ Items (Expected Relationship to Hypotheses Based on Visual Inspection)

Item No.	Text of CMAQ Item	Direc -tion
Index 1. *Hypothesis: Highly reliable teams maintain open and flexible communication and are able to shift responsibility as needed, based on a shared mental model of the situation.*		
5	The pilot flying the aircraft should verbalize plans for procedures or maneuvers and should be sure that the information is understood and acknowledged by affected crew members.	+
7	Pilots-in-command should encourage pilots and crew chiefs to question procedures and flight profile deviations during normal flight operations and in emergencies.	+
15	The pilot-in-command is solely responsible for the leadership of the crew team.	-
17	When joining a unit, a new crew member should not offer suggestions or opinions unless asked.	-
19	Pilots-in-command who accept and implement suggestions from the crew lessen their stature and reduce their authority.	-
20	Crewmembers should monitor the pilot-in-command's performance for possible mistakes and errors.	+
23	The pilot-in-command should seek advice from crew members when updating mission plans.	+
24	The pilot-in-command should use his crew to help him maintain situation awareness.	+
25	The pilot-in-command is solely responsible for maintaining awareness of crew capabilities.	-
26	Only when the pilot-in-command is overloaded should he pass responsibility to other crew members.	-
37	Nonrated crew members should be actively involved in planning the mission.	+
38	Understanding the commander's concept is of minor importance to mission execution.	-
40	Thinking through difficult segments, events, and tasks is primarily the pilot-in-command's responsibility.	-
46	External circumstances require crew members to provide situational leadership for short periods of time.	+

Index 2. *Hypothesis: Highly reliable crews are sensitive to each other's performance and workload, and have mutual mental models that allow them to anticipate each other's needs.*		
2	Crew members should monitor each other for signs of stress or fatigue and should discuss the situation with the affected crew member(s).	+
11	Effective crew coordination requires crew members to take into account the personalities of other crew members.	+
27	Crewmembers should be aware of other crew member's workload.	+
28	If a crew member is having difficulties executing his responsibilities, other crew members should provide assistance.	+
39	Each crew member should watch for situations in which external events limit others' performance.	+
44	Crew members should be able to anticipate requirements as the mission progresses.	+

we combined into a second index by taking an average. These two CMAQ indices are independent—their correlation is .04 before training and -.15 after training.

Tables 11 and 12 show the correlation of the two CMAQ indices with the communications measures pre-training and post-training. The CMAQ indices are much more strongly related to the communication measures *before* training (12 significant correlations) than after training (five significant correlations). We suggest that the attitudes of the individual crew members were a strong driver of crew communication and coordination patterns before training, whereas after training the crew's communication and coordination patterns were more influenced by the *training itself* than by the attitudes of the crew members. There was also less variability in the CMAQ indices after training. The standard deviation for Index 1 was .318 before training and .305 after training, a small decrease of about 4 percent. The standard deviation for Index 2 showed a much larger decrease—from .368 to .250 or 32 percent. Apparently one effect of the training was to produce more homogeneous attitudes among the crews concerning the importance of cross-monitoring behavior.

We expected that the overall communication rate and the type of communication would be related to Index 1, which combines CMAQ items concerning open two-way communication in the crew and the flexibility of responsibilities in response to changes in the situation. We expected that Index 2, which combined CMAQ items dealing with cross monitoring and sharing of workload among the crew members, would be positively related to anticipation ratios.

Table 11

Pre-Training Correlation of CMAQ Indices with Measures

Communication Measures	CMAQ Index 1 Open and Flexible Communication	CMAQ Index 2 Cross-Monitoring & Anticipation
Index Mean	4.04	5.87
Index Standard Deviation	.318	.368
ANTICIPATION RATIOS	Correlation	Correlation
Total Anticipation Ratio	-.09	.36‡
Anticipation Ratio for Information Transmissions	-.47*	-.06
Anticipation Ratio for Actions and Tasks	.13	.36‡
Anticipation Ratio in Routine Conditions	-.02	.37‡
Anticipation Ratio in Crisis Conditions	-.24	.10
COMMUNICATION RATES		
Communication Rate—Total	.24	.01
Proportion of Communication Rate for Information Transmissions	-.38‡	.32
Proportion of Communication Rate for Actions and Tasks	.10	-.42‡
Proportion of Communication Rate for Planning and Problem Solving	.42‡	.08
Communication Rate in Routine Conditions	.23	-.03
Communication Rate in Crisis Conditions	.33	.29
Proportion of Communication Rate for Information Transmissions in Routine Conditions	-.36‡	.27
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	.10	-.35‡
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	.40‡	.07
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	-.30	.38‡
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	.11	-.47*
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	.26	.05

Table 12

Post-Training Correlation of CMAQ Indices with Communication Measures

Communication Measures	CMAQ Index 1 Open and Flexible Communication	CMAQ Index 2 Cross-Monitoring & Anticipation
Index Mean	4.18	6.10
Index Standard Deviation	.305	.250
ANTICIPATION RATIOS	Correlation	Correlation
Total Anticipation Ratio	-.19	-.09
Anticipation Ratio for Information Transmissions	-.30	-.10
Anticipation Ratio for Actions and Tasks	-.02	-.08
Anticipation Ratio in Routine Conditions	-.14	-.10
Anticipation Ratio in Crisis Conditions	-.53*	-.12
COMMUNICATION RATES		
Communication Rate—Total	-.12	.05
Proportion of Communication Rate for Information Transmissions	.18	-.46*
Proportion of Communication Rate for Actions and Tasks	.01	.33
Proportion of Communication Rate for Planning and Problem Solving	-.25	.23
Communication Rate in Routine Conditions	-.16	-.00
Communication Rate in Crisis Conditions	.06	.21
Proportion of Communication Rate for Information Transmissions in Routine Conditions	.16	-.42‡
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	.01	.31
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	-.24	.20
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	.10	-.53*
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	.08	.38‡
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	-.24	.33

Table 11 shows that these hypotheses were partially supported by the pre-training data. While there was no significant correlation between Index 1 and the volume of communication in routine or crisis conditions, there was a relationship between Index 1 and the type of communications observed. Crews that ranked the items in Index 1 as more important produced fewer communications that transferred information and more that concerned planning and problem solving, especially under routine conditions. There was a negative correlation between Index 1 and the anticipation ratio for information transmission, related to the finding that team with a higher score on Index 1 produced fewer communications dealing with information transfer.

Index 2 was, as expected, correlated with anticipation ratios as shown in Table 11. There was a significant positive correlation for the total anticipation ratio, the anticipation ratio for actions and tasks, and the anticipation ratio in routine conditions. Crews that rated cross monitoring functions as more important were more likely to anticipate the needs of other crew members. There was also a correlation between Index 2 and the type of communication. Crews with higher scores on Index 2 produced proportionately fewer communications dealing with actions and tasks in both routine and crisis conditions. This result is consistent with the finding that Index 2 was positively correlated with the anticipation ratio for actions and tasks. Teams that rated cross-monitoring as important appear to have anticipated the need for certain actions and tasks, reducing the number of requests for these actions and, therefore, the proportion of communications dealing with actions and tasks.

The post-training correlations in Table 12 show only one significant correlation between Index 1 and the communication measures—a negative correlation with the anticipation ratio during crisis conditions. This finding is difficult to explain and is possibly an anomaly. After training, crews with a higher rating on Index 2 produced proportionately fewer communications concerned with information transmission in both routine and crisis conditions, and proportionately more communications dealing with actions and tasks during crisis conditions. Crews in general produced proportionately fewer communications involving information transmission after training (see Table 7). This effect seems to have been especially pronounced for the crews that placed more importance on cross-monitoring.

Relationship of Communication and Coordination Patterns to Crew Performance

A major goal for the initial testing of the communication-analysis instrument was to determine if communication measures based on this instrument could be related to crew performance. This subsection analyzes the relationship between communication measures and crew performance from several different perspectives using a variety of different performance measures.

There was no single universal measure of the performance of the helicopter crews. Instead, a variety of measures were available from the DRC evaluation effort, including an overall grade provided by the IPs, ratings on 18 ATM tasks provided by the IPs, the percentage of mission sub-objectives completed, and ratings on six measures of mission performance developed by DRC.[7] We computed several overall composite scores based on these measures, including an average ATM score[8] and a total mission-performance score based on the six mission-performance measures. Tables 13 and 14 show the correlations of these performance measures with the 17 communication measures derived from the videotapes before training (Table 13) and after training (Table 14).

Anticipation ratios and performance. We expected that anticipation ratios would be positively correlated with crew performance, especially in crisis conditions, where we expect the teams to rely on their mutual mental models to anticipate the needs of the other crew members. Before training, anticipation ratios were correlated with three of the six mission performance measures—terrain flight navigation, response to aircraft emergencies, and mission-threatening crew errors. A higher anticipation ratio for information exchange was associated with fewer flight deviations. A higher anticipation ratio for actions and tasks was associated with better performance in aircraft emergencies. A higher anticipation ratio for information exchange and a higher anticipation ratio in crisis conditions were associated with fewer mission-threatening crew errors and a better response to those errors. After training, anticipation ratios were even more strongly correlated with better terrain flight navigation.

In the pre-training data, anticipation ratios were highly correlated with the percentage of mission objectives completed, and the anticipation ratio in crisis conditions was associated with a higher average ATM score. After training, only the anticipation ratio for actions and tasks was correlated with the percentage of objectives completed, but the total anticipation

---

[7] We selected representative measures within each of the six mission-performance areas established by DRC. The measures selected are documented in Appendix A. The six mission-performance ratings have been adjusted so that a higher score always indicates better crew performance.

[8] We also computed the correlations between each of the 18 individual ATM measures and our 17 communication measures pre- and post-training. These correlations are presented in Appendix D. The communication and coordination measures were significantly related to some individual ATM tasks but not to others. Pre- and post-training correlation patterns are quite different, and it is difficult to interpret the results without a more intensive and detailed task analysis of the components comprising each of the 18 ATM measures.

Table 13

Correlation of Communication Measures with Mission-Performance Measures Before Training

Communication Measures	PERF1 Terrain flight navig. (dev.)	PERF2 Threat avoid. & evas.	PERF3 Aircraft emergencies	PERF4 Unexpected events	PERF5 Instr. flight recov.	PERF6 Mission threat crew error	% objectives completed	Total performance (1-6)	Avg. ATM Score	Overall grade (subjective)	Number of signif correlations
	Performance Measures Pre-Training										
ANTICIPATION RATIOS											
Total Anticipation Ratio	.05	-.07	.18	.07	.08	.31	.58*	.02	.25	.00	1
Anticipation Ratio for Information Transmissions	.42‡	-.02	-.14	.09	.09	.71**	.28	-.01	-.01	-.18	2
Anticipation Ratio for Actions and Tasks	-.15	-.22	.37‡	.06	-.04	.09	.45*	-.11	.11	.06	2
Anticipation Ratio in Routine Conditions	.03	-.14	.16	.01	.09	.19	.50*	-.07	.19	-.04	1
Anticipation Ratio in Crisis Conditions	.28	.17	.25	.23	.05	.69**	.74**	.25	.37‡	.16	3
COMMUNICATION RATES											
Communication Rate—Total	-.29	.17	.30	-.60**	-.42‡	-.35‡	.08	.06	.06	.10	3
Proportion of Communication Rate for Information Transmissions	.36‡	-.01	.03	.18	.20	.28	-.09	-.01	.42‡	.12	2
Proportion of Communication Rate for Actions and Tasks	-.33	-.26	.07	-.07	-.31	.00	-.26	-.22	-.57*	-.06	1
Proportion of Communication Rate for Planning and Problem Solving	-.11	.33	-.13	-.16	.10	-.39‡	.45*	.29	.12	-.09	2
Communication Rate in Routine Conditions	-.28	.17	.30	-.58*	-.41‡	-.35‡	.06	.07	-.11	.05	3
Communication Rate in Crisis Conditions	-.24	.07	.23	-.43‡	-.37‡	-.22	.11	.00	.33	.40‡	3
Proportion of Communication Rate for Information Transmissions in Routine Conditions	.29	-.09	.04	.13	.13	.25	-.14	-.10	.41‡	.18	1
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	-.30	-.17	.06	-.03	-.31	.00	-.24	-.15	-.61**	-.13	1
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	-.03	.36‡	-.15	-.15	.22	-.35‡	.51*	.35‡	.20	-.09	4
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	.52*	.39‡	-.10	.33	.43‡	.30	.20	.39‡	.27	-.21	4
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	-.27	-.47*	.16	-.30	-.13	-.05	-.19	.42‡	-.06	.29	2
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	-.36‡	.03	-.06	-.09	-.41‡	-.32	-.04	-.03	-.27	-.06	2
Number of Significant Correlations	4	3	1	3	5	6	6	3	5	1	37

Table 14

Correlation of Communication Measures with Mission-Performance Measures After Training

Communication Measures	PERF1 Terrain flight navig. (dev.)	PERF2 Threat avoid. & evas.	PERF3 Aircraft emergencies	PERF4 Unexpected events	PERF5 Instr. flight recov.	PERF6 Mision threat crew error	% objectives completed	Total performance (1-6)	Avg. ATM Score	Overall grade (subjective)	Number of signif correlations.
	Performance Measures Post-Training										
ANTICIPATION RATIOS											
Total Anticipation Ratio	.56*	.02	.25		.04	.07	.17	-.10	.41‡	.44*	3
Anticipation Ratio for Information Transmissions	.33	.11	-.15		.05	.18	.10	.03	.13	.38‡	1
Anticipation Ratio for Actions and Tasks	.49*	.19	.32		-.18	-.09	.39‡	-.04	.28	.17	2
Anticipation Ratio in Routine Conditions	.56*	.07	.22		.00	.06	.16	-.08	.38‡	.37‡	3
Anticipation Ratio in Crisis Conditions	.39‡	-.16	.31		.20	.15	.17	-.12	.34	.57*	2
COMMUNICATION RATES											
Communication Rate—Total	.42‡	.03	-.28		-.24	-.03	.35‡	-.22	.17	.39‡	3
Proportion of Communication Rate for Information Transmissions	.33	.00	.62**		.09	-.27	-.25	-.07	.07	.00	1
Proportion of Communication Rate for Actions and Tasks	-.20	.23	-.46*		-.26	.02	.28	.12	-.26	.22	1
Proportion of Communication Rate for Planning and Problem Solving	-.21	-.28	-.29		.18	.35‡	.00	-.05	.22	.26	1
Communication Rate in Routine Conditions	.45*	.01	-.28		-.22	-.03	.32	-.24	.14	.39‡	2
Communication Rate in Crisis Conditions	.18	.06	-.26		-.31	-.10	.39‡	-.16	.42‡	.35‡	3
Proportion of Communication Rate for Information Transmissions in Routine Conditions	.42‡	.05	.59*		.15	-.19	-.26	-.02	.09	.07	2
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	-.30	.17	-.43‡		-.31	-.04	.31	.08	-.27	-.28	1
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	-.22	-.30	-.30		.18	.31	-.02	-.08	.20	.26	0
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	-.11	-.22	.42‡		-.17	-.50*	-.09	-.28	-.10	-.26	2
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	.24	.42‡	-.35‡		.08	.23	.03	.32	-.07	.10	2
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	-.12	-.17	-.22		.17	.47*	.09	.05	.24	.26	1
Number of Significant Correlations	7	1	6		0	3	3	0	3	7	30

*** No Variance

ratio was correlated with the average ATM score, as well as with the overall grade for the crew.

Overall, we conclude that anticipatory behavior, as measured by the anticipation ratio, was positively related to crew performance both prior to and after training. This relationship was especially strong for terrain flight navigation, where anticipatory behavior seems to have reduced deviations from the flight path. We had expected that anticipatory behavior would be more strongly related to performance during crises than during routine conditions. We cannot test this hypothesis directly from the data available because the overall performance data are for the entire scenario, not just for the crisis periods. We do note that the anticipation ratio for actions was correlated with better performance in emergencies prior to training, and positively, but not significantly, related to better performance in emergencies after training. Also, the correlation between anticipation ratio and overall grade after training is stronger for anticipation during crisis conditions than for anticipation during routine conditions.

Communication rates and performance. We did not necessarily expect that a higher overall communication rate would be associated with better crew performance. More communication in a crew is not necessarily better—it depends on the nature and circumstances of that communication. Prior to crew training we found negative relationships between overall communication rate and response to unexpected events, instrument flight recovery, and mission-threatening errors. There was no correlation between overall communication rate and any of the overall performance measures, although the communication rate in crisis conditions was associated with a higher overall grade for the crew.

After training, the communication rate shows a more positive pattern. Higher overall communication rates were associated with better terrain flight navigation, a higher percentage of objectives completed, and a higher overall grade. Higher communication rates during routine conditions were associated with better terrain navigation and with a higher overall grade. Higher communication rates during crisis conditions were associated with a greater percentage of objectives completed, higher average ATM scores, and a higher overall grade. We conclude that not only was there a higher volume of communication after training, but this more-frequent communication was of a type that improved performance.

Communication functions and performance. The relationship between the proportion of communications devoted to each of three functions (information exchange, actions and tasks, or planning and problem solving) and performance is not straightforward. Before training, a higher proportion of communications dealing with information exchange, and a lower proportion dealing with actions and tasks, was associated with better terrain navigation

and a higher average ATM score. A higher proportion dealing with planning and problem solving was associated with a higher percentage of objectives completed.

After training, the proportion of communication by function was not associated with any of the overall performance measures (i.e., average ATM score, percentage of objectives completed, or grade). One pattern noted in the post-training data (but not in the pre-training data) was that better performance in aircraft emergencies was associated with a higher proportion of communication dealing with information transfer and a lower proportion of communication dealing with other functions. We suggest that well-trained crews know what to *do* in an emergency—all they need is an accurate assessment of the situation that is provided by the exchange of information. After training, however, a higher proportion of communications dealing with planning and problem solving, both overall and during crisis conditions, is associated with a lower incidence of mission-threatening crew errors, suggesting that there is a complex relationship between the nature of the communication patterns and performance in emergency situations.

Comparison of lowest- and highest-performance crews. As an alternative method to gain insight into the communication patterns of high-performance crews, we selected the three highest-performance crews and the three lowest-performance crews both before and after training, and compared their scores on the 17 communication measures. We selected the low- and high-performance groups by rank ordering the 15 crews based on their overall grade. To achieve a good separation of the two groups, we selected the top 20 percent (three crews) and the bottom 20 percent (three crews) of the crews. The remaining 60 percent of the crews were considered intermediate in performance. When two or more crews had the same overall grade, we used the average ATM score to break the tie. Groups were selected separately based on the pre-training data and the post-training data.

The communication measures for the three lowest- and highest-performance crews after training are shown in Table 15. The comparison of the pre-training groups did not reveal any significant differences, and these results are not presented. Table 15 shows three significant differences between the lowest- and highest-performance crews. The anticipation ratios were all higher for the highest-performance group. This difference was significant for the overall anticipation ratio, and for the anticipation ratio in crisis conditions. Communication rates were higher for the high-performance crews, but this difference was not significant. In crisis conditions the proportion of communications dealing with the exchange of information was significantly lower for the highest-performance crews. The proportion of communications dealing with planning and problem solving appeared higher for the highest-performance crews, but the difference was not significant.

Table 15

Communication Measures for Three Lowest- and Highest-Performance Crews After Training

Communication Measures	Lowest-Performance Crews	Highest Performance Crews	t-test for Contrast[1]
ANTICIPATION RATIOS	Mean	Mean	
Total Anticipation Ratio	2.47	3.18	1.75‡
Anticipation Ratio for Information Transmissions	4.14	4.80	.93
Anticipation Ratio for Actions and Tasks	1.16	1.42	-.81
Anticipation Ratio in Routine Conditions	2.41	3.06	1.53
Anticipation Ratio in Crisis Conditions	2.81	3.72	2.05*
COMMUNICATION RATES			
Communication Rate—Total	11.28	14.13	1.26
Proportion of Communication Rate for Information Transmissions	.52	.52	.00
Proportion of Communication Rate for Actions and Tasks	.40	.36	-1.24
Proportion of Communication Rate for Planning and Problem Solving	.08	.12	1.24
Communication Rate in Routine Conditions	11.27	14.03	1.18
Communication Rate in Crisis Conditions	11.29	14.62	1.39
Proportion of Communication Rate for Information Transmissions in Routine Conditions	.50	.52	.44
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	.42	.36	-1.52
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	.08	.12	1.24
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	.60	.52	-1.76‡
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	.33	.36	.76
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	.07	.12	1.55

These findings confirm the pattern noted from the correlations. A higher anticipation ratio was associated with better crew performance, and there was some indication that anticipatory behavior was especially important in crisis conditions. A higher communication rate *per se* did not necessarily imply better performance, but higher-performance crews seemed to communicate more. Higher-performance crews also seemed to differ from lower-performance crews in the relatively lower proportion of their communication that was devoted to information exchange versus planning and problem solving under crisis conditions. The nature of this difference (relative, not absolute) has not yet been fully explored.

Conclusions regarding communication and performance. We conclude that overall anticipation ratios have a relatively robust relationship to overall crew performance. Better crews anticipate each other's needs. There is some evidence that anticipation ratios during crisis conditions may be related to performance during crisis conditions, but this hypothesis cannot be tested directly with the data available. Higher overall communication rates were associated with higher overall crew performance after but not before training, indicating that the nature of the communication as well as its volume affects performance. Communications by function do not have a straightforward relationship to overall crew performance, and the findings concerning communication by functions in crisis situations are inconclusive. Measures that are more precisely focused on performance during crisis periods are needed in order to explore the relationship between crew communication and coordination patterns and crew performance in a crisis.

Conclusions and Recommendations

We had two major objectives for the observation and analysis of the communication patterns for helicopter flightcrews in the Fort Campbell videotapes. First, we wanted to test the feasibility and reliability of our measurement instrument and methodology. Second, we wanted to test the validity of our measurement approach by comparing our communication-measurement results with other teamwork measures and by assessing the extent to which the results obtained are consistent with the predictions of theories about team coordination. We present conclusions on each of these issues below.

Measurement Instrument and Methodology

Overall, we were satisfied with the feasibility of the measurement approach and the reliability of the communication-analysis instrument. The coefficient alpha statistics for inter-rater reliability were quite high for the individual measures collected by the two raters who observed the videotapes. We conclude that the raters were able to reliably record communications by type (requests and transfers) and function

(information transmission, actions and tasks, and planning and problem solving). We were not successful, however, in having our raters, who were not domain experts, evaluate the *quality* of the coordination of the crews being observed. The reliability was low for both of the coordination-evaluation instruments tested. We conclude that ratings of coordination quality must be provided by domain experts who have a detailed understanding of the activities being observed.

An important limitation of the approach used was that we were not able to reliably identify, based solely on the videotapes, the hierarchical roles being played by the two individuals being observed. Two designators for the individuals in the flightcrew may be of interest—the individual who is in command, and the individual who is actually piloting the aircraft. The first is invariant during a flight, while the second can change several times. We were able to reliably distinguish utterances produced by each of the two individuals in the crews based on the videotapes, but we could not always determine which was the commander from the videotapes, and we were not able to obtain that information from the Testbed records. We could usually distinguish which individual was flying the aircraft from the nature of the communications, but we had not designed the instrument to record this information.[9] For future applications we will need, at a minimum, to obtain information about which individual is the commander, and we may want also to collect information about which individual is acting as the pilot at any given point in the scenario.

Suggested modifications to the coding instrument. Based on our initial experience using the communication-analysis instrument and the results of our analysis of the data, we have identified several ways in which the coding procedure can be refined and improved for future applications. The major improvements concern the rules for switching coding sheets and the categories used for classification.

The raters used a coding sheet for every 15-minute segment of the Fort Campbell videotapes. The major reason for segmenting the coding in this way was so that if the coders disagreed (i.e., inter-rater reliability was low) they could localize the disagreement to a particular segment on the tape. Now that we have established that the reliability of the instrument is high, we suggest that it may be more meaningful to tie the actions on the recording forms to the role that the person is playing. In the simulations we observed, the two individuals sometimes switched responsibility for flying the plane. In some instances

[9] In the videotape used for testing and calibration, and in the first few videotapes coded, the first individual to fly the aircraft always served as the pilot. We did not, therefore, anticipate that the roles would shift in the other videotapes and did not design the form to capture this information.

this change in roles lasted for a very short period (about 10 seconds), whereas in other instances it was a more enduring change. To capture these changes, the raters could start a new form each time there is a switch in roles. In that way we could link the nature of the communications to the individual's role. For role changes that last only a brief time (sometimes only a few seconds), it may be sufficient to note the change on bottom of the coding sheet, since the number of interchanges that can occur in such a short period will be minimal. The adaptive changing of roles is predicted by the theory of adaptive team behavior, and recording it will enable us to test hypotheses about the crew's flexibility in restructuring their responsibilities in response to changes during the mission.

We may also want to switch from one coding sheet to another based on the occurrence of pre-specified events rather than at equal time intervals. For example, the scenarios we observed were divided into five segments, and it would be possible for raters to tie the coding sheets to the mission segments. This could be useful when there were hypotheses relating the communication patterns to different types of situations.

A second change for subsequent analyses concerns the transfer action/task category classification. In a subsequent version, we may wish to discriminate acknowledgments, such as an "O.K." which directly follows a request or statement, implying "I am doing what you asked" or "I acknowledge what you said," from statements about the performance of non-requested actions. Following a suggestion made by Orasanu (1990), if one of the crew members reads off a check list, we suggest that in the future this be coded as one utterance, rather than coding each item on the check list as an action completed.

There is some ambiguity about whether statements of actions to be done in the future represent transfers of action or transfers of planning.[10] There is a certain arbitrariness about when an impending action becomes planning. Raters can fairly easily agree upon a convention and thereby maintain a reliable coding system, but there is a question of validity in capturing the speaker's intention. There is a similar question about when a transfer of information becomes planning. For example, is "We should see a road coming from left to right" a transfer of information (as it was coded), or a way of planning for the future? The interpretation as planning implies a future time frame. In some cases, the intention is not completely clear.

[10] For example, "I will call your turns" or "I will start a fuel check as soon as we get in the air." For the Fort Campbell tapes the raters classified these as actions, but one could also view them as transfers of plans. The ambiguity is particularly clear in the second example. If the statement had been "we'll start a fuel check when we get to ..." it would be coded as planning.

Our results indicate the planning/problem-solving function is an important factor in effective cockpit communication. In coding the Fort Campbell tapes, when a problem-solving session was initiated we used a single mark to indicate problem solving, and then recorded the subsequent discourse using more-specific categories. This method does not capture the duration or the volume of the problem-solving discourse. In the future it might be useful to have a way to capture that information.

The raters found very few instances where utterances fell into the "request for planning" category. After coding the tapes, however, we have some examples of utterances in that category and have a clearer idea of category membership. We suggest retaining the category at least until we code another set of tapes.

Even though the raters were able to code directly from the tapes, and could sometimes work in real time, the coding of the tapes requires a significant block of time and effort. In order to minimize the amount of time required for coding the tapes, in the future we would consider sampling the tape instead of coding the entire tape. Certain "significant" portions of the mission (in particular, crisis situations) could be completely analyzed, wile other more routine portions of the mission could be sampled.

Measures and Theory

Our second major concern was the validity of the communication measures obtained from the videotapes. Validity may be assessed from several different perspectives. One indication of validity is sensitivity to manipulations that are expected to affect communication: did the measures show changes between pre- and post-training behavior? Another indication of validity is correlation to other measures of teamwork and to measures of crew performance. A final indication of validity is the presence of patterns in the data that are consistent with expectations based on a model of crew coordination. We assess the results for each of these indications in turn.

Sensitivity to the effects of training. Several of the communication measures showed differences between pre- and post-training behavior. The anticipation ratio was higher after training, especially for actions and tasks, and in routine conditions. Training increased the communication rate from about 10 communications per minute before training to almost 12 per minute after training. The nature of the communication changed as well, and the effects of training on the nature of the communication were different in routine and crisis conditions. In routine conditions there was proportionately less exchange of information and more planning and problem solving after training. In crisis conditions there was proportionately less communication about actions and tasks after training.

Correlation with other teamwork measures. We examined the correlation of the communication measures with selected basic

qualities from the ACE and with two indices computed from the CMAQ questionnaire items. After training there was considerable correlation of the communication measures with the ACE basic qualities related to teamwork. There was no one-to-one mapping between our measures and the ACE qualities, however. Typically, a number of communication measures were correlated with each ACE dimension, and a number of ACE dimensions with each of our measures. All of the significant correlations were positive, i.e., higher values for the communication measures were associated with higher scores on the ACE dimensions (better teamwork). Specifically, anticipation ratios, communication rates, and the proportion of communication that involved planning and problem solving were positively related to a number of ACE basic qualities. We conclude that the communication measures capture behaviors that are associated with better teamwork, but in a way that cuts across many of the ACE dimensions.

Results for selected CMAQ items also suggest that the communication measures are capturing elements of teamwork. Correlations between indices based on the CMAQ and the communications measures were stronger before training, suggesting that individual crewmember attitudes were more heterogeneous and a stronger predictor of behavior before the crews were trained. Before training, an index combining CMAQ items concerning open and flexible communication and distribution of responsibilities was positively associated with more planning and problem solving and less exchange of information in routine conditions. A second index that combined items related to cross-monitoring and anticipation of needs was positively related to anticipation ratios in both routine and crisis conditions. We found that the crew's attitudes on these two indices were much more strongly correlated to their behavior before than after the crew-coordination training, however. We conclude that the attitudes toward teamwork captured by the CMAQ are related to the behaviors captured by our communication measures.

Correlation with performance measures. We expected that some communication patterns would be associated with better crew performance, and we hoped that our measures would be sensitive to performance differences. After crew training, we found that anticipation ratios were positively associated with overall performance, as were communication rates. Results are less clear for communication by function, indicating that the relationship between the nature of the communications observed and the performance of the crew should be captured on a task-by-task basis rather than by overall performance measures.

Theory of crew coordination. An important indication of the validity of the communication measures is their congruence with a theory of crew coordination. Do the results "make sense" from the perspective of the theory? The theory suggests that crews have mutual mental models of the other crew members, as well as a shared mental model of the situation, and that these mental models allow them to anticipate the needs of the other crew members.

Further, we expect that this anticipatory behavior will be more pronounced under high-workload conditions, and that the presence of this implicit coordination will be associated with better performance under high workload.

The results from the Fort Campbell observation and analysis are relevant to part, but not all, of this theory. We observed anticipatory behavior, as measured by anticipation ratios. This anticipatory behavior increased after training, and was associated with better crew performance according to a number of indicators. We had no explicit measure of workload in the analysis but, if we assume that workload was greater in crisis conditions (as they were defined in the simulation scenarios), we find that anticipation ratios were higher in crisis than in routine conditions, indicating that implicit coordination was being used as predicted by the crew-adaptation premise.

We were not able to directly test the association between the crews' use of implicit coordination and performance levels under high-workload conditions, however. We did not have workload measures, and we did not have measures of performance that were well-linked to the conditions in which high workload occurred. Based on overall measures of performance, we did find that after training the anticipation ratio during crisis conditions was more strongly related to the crew's overall grade than was the anticipation ratio during routine conditions. More finely focused measures of crew performance during high workload are needed to fully test this part of the theory.

The results of the analysis provide some insight into how crews change their communication patterns in emergency conditions. Before training there was little difference in the nature of the communications during routine and crisis conditions. After training, crews produced proportionately more communications involving information exchange during a crisis, and less communications involving actions and tasks. We suggest that the well-trained crew knows what to do in an emergency, so that little explicit communication regarding actions and tasks is required. We also found that a higher proportion of communication involving information exchange, and a lower proportion involving actions and tasks, was associated with *better* emergency performance after training.

Overall, we conclude that the communication measures tested show reasonable validity from a number of different perspectives. They are sensitive to the effects of training; they are correlated with, but not identical to, other teamwork measures; they are sensitive to differences in crew performance; and they provide partial support for the theory of crew coordination. Workload measures, and more universally accepted measures of mission performance in high-workload conditions, are needed for a more complete assessment of the theory.

References

Baker, D.P. & Salas, E. (1992). Principles for measuring teamwork skills. Human Factors, 34 (4), 469-475.

Barrett, L.E. (1993). Decision making teams: Their study in the U.S. military (Technical Report No. 93-1). East Lansing, MI: Michigan State University.

Cannon-Bowers, J.A., Salas, E., & Converse, S. (1990). Cognitive psychology and team training: Training shared mental models of complex systems. Human Factors Society Bulletin, 33 (12), 1-4.

Carley, K.M. (1990). Coordinating for success: Trading information redundancy for task simplicity. In Proceedings of the 23rd Hawaii International Conference on Systems Science. Honolulu, HW.

Cronbach, L. J. (1970). Essentials of psychological testing. New York: Harper and Row.

Diehl, A. (1991). The effectiveness of training programs for preventing aircrew "error." In Proceedings of the Sixth International Symposium on Aviation Psychology (pp. 640-655). Columbus, OH: The Ohio State University.

Dwyer (1992). AAW Team Observation Measures (ATOM), private communication.

Foushee, H.C. & Helmreich, R. L. (1988). Group interaction and flight crew performance. In E.L. Wiener & D.C. Nagel (Eds.), Human factors in aviation. San Diego: Academic Press.

Geis, C.E. (1987). Changing attitudes through training: A formal evaluation of training effectiveness. In R.H. Jensen (Ed.), Proceedings of the 4th International Symposium on Aviation Psychology (pp. 392-398). Columbus, OH: The Ohio State University.

Green, R. (1990). Human error on the flight deck. In D.E. Broadbent, J. Reason, & A. Baddely (Eds.), Human factors in hazardous situations. Oxford: Clarendon Press.

Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human mental workload. Amsterdam: Elsevier.

Helmreich , R.L. (1984). Cockpit management attitudes. Human Factors, 26, 583-589.

Helmreich, R.L. & Wilhelm, J.A. (1986). Evaluating CRM training: I. Measures and methodology (NASA/UT Technical Report No. 86-8). Moffett Field, CA: NASA Ames Research Center.

Helmreich, R.L. & Wilhelm, J.A. (1987). Evaluating cockpit resource management training. In R. S. Jensen (Ed.), Proceedings of the Fourth International Symposium on Aviation Psychology (pp. 440-446). Columbus, OH: The Ohio State University.

Kalisetty, S., Kleinman, D.L., Serfaty, D., & Entin, E.E. (1993). Coordination in hierarchical information processing structures (CHIPS). In Proceedings of the 1993 JDL Command and Control Research Symposium. McLean, VA: SAIC.

Kleinman, D.L., Pattipati, K.R., Luh, P.B., & Serfaty, D. (1992). Mathematical models of team decision making. In R.W. Swezey & E.S. Salas (Eds.), Teams: Their training and performance. Norwood, NJ: Ablex.

Kleinman, D.L. & Serfaty, D. (1989). Team performance assessment in distributed decisionmaking. In Gibson et al. (Eds.), Proceedings of the Symposium on Interactive Networked Simulation for Training. Orlando, FL.

Lautman, L.G. & Gallimore, P.L. (1987). Control of the crew caused accident: Results of a 12-operator survey. Airliner. Seattle, WA: Boeing Commercial Airplane Co., 1-6.

LaPorte, T.R. & Consolini, P.M. (1988). Working in practice but not in theory: Theoretical challenges of high reliability organizations. Annual Meeting of the American Political Science Association. Washington, DC.

MacIntyre, R.M., Morgan, B.B., Jr., Salas, E.S., & Glickman, A.S. (1988). Teamwork from team training: New evidence for the development of teamwork skills during operational training. Paper presented at the Interservice/Industry Training Systems Conference, Orlando, FL.

Meister, D. & Enderwick, T.P. (1992). Special issue preface, Human Factors, 34(4), 383-385.

Nunnally, J. (1967). Psychometric Theory. New York: McGraw Hill.

Orasanu, J.M. (1990). Shared mental models and crew decision making (CSL Report 46). Princeton, NJ: Princeton University, Cognitive Science Laboratory.

Orasanu, J.M. & Fischer, U. (1992). Team cognition in the cockpit: Linguistic control of shared problem solving. In Proceedings of the 14th Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Erlbaum.

Oser, R., McCallum, G.A., Salas, E., & Morgan, B.B. Jr. (1989). Toward a definition of teamwork: An analysis of critical team behaviors (TR-89-004). Orlando, FL: Human Factors Division, Naval Training Systems Center.

Pfeiffer, J. (1989, July). The secret of life at the limits: cogs become big wheels, Smithsonian.

Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. In D.E. Broadbent, J. Reason, & A. Baddely (Eds.), Human factors in hazardous situations. Oxford: Clarendon Press.

Reid, G.B., Singledecker, C.A., Nygren, T.E. & Eggemeir, F.T. (1981). Development of multidimensional subjective measures of workload. In Proceedings of the 1981 IEEE International Conference on Cybernetics and Society (pp. 403-406). Atlanta, GA.

Salas, E.S., Dickinson, T., Converse, S.A. & Tannenbaum, S. (1992). Toward an understanding of team performance and training. In R.W. Swezey & E.S. Salas (Eds.), Teams: Their training and performance. Norwood, NJ: Ablex.

Serfaty, D., Entin, E.E., & Deckert, J.C. (1993). Team adaptation to stress in decision-making and coordination with implications for CIC team training (TR-564). Burlington, MA: ALPHATECH, Inc.

Serfaty, D., Entin, E.E., & Volpe, K. (1993a). Adaptation to stress in team decision making and coordination. In Proceedings of the 37th Annual Meeting of the Human Factors and Ergonomics Society (pp. 1228-1232). Santa Monica, CA.

Serfaty, D., Entin, E.E., & Volpe, K. (1993b). Implicit coordination in command teams. In Proceedings of the 1993 JDL Command and Control Research Symposium. McLean, VA: SAIC.

Simon, R. (1991). Technical report: Results of the data analysis Army aircrew coordination measures testbed conducted Spring 1990. Wilmington, MA: Dynamics Research Corporation.

Simon, R. & Grubb, G. (1993). Validation of crew coordination training and evaluation methods for Army aviation (E-785u). Wilmington, MA: Dynamics Research Corporation.

Simon, R. & Grubb, G. (1992). Development of crew coordination evaluation methods and materials. Wilmington, MA: Dynamics Research Corporation.

Stout, R., Cannon-Bowers, J., Morgan, B.B., & Salas, E. (1989). The development of a scale to assess the teamwork needs of training situations. In Proceedings of the Human Factors Society 33rd Annual Meeting. Santa Monica, CA.

Swezey, R.W. & Salas, E. (1992). Guidelines for use in team training development. In R.W. Swezey & E.S. Salas (Eds.), Teams: Their training and performance. Norwood, NJ: Ablex.

Vaughan, W. (1990). A theory of the effects of advanced information technologies on organizational design, intelligence, and decision making. In Proceedings of the 1990 Symposium on Command and Control Research. McLean, VA: SAIC.

Wagner, M., Simon, R. & Leedom, D. (1990). Development of measures of crew coordination. Wilmington, MA: Dynamics Research Corporation.

Appendix A

Questionnaires

CREW MEMBER PRE-MISSION QUESTIONNAIRE

CREW ID: ____________________	CREW POSITION: Pilot / CPG
DATE/TIME: __________________	LAST 4 DIGITS OF SS#: _________

1. Briefly list up to three potential "show stoppers" that could compromise this mission.

   1)____

   2)____

   3)____

2. Briefly list up to three of the most critical requirements for good crew coordination during this mission.

   1)____

   2)____

   3)____

3. Briefly describe the two most important tasks/responsibilities for you while at the battle position (BP).

   1)____

   2)____

4. Briefly describe the two most important tasks/responsibilities for your fellow crew member while at the BP.

   1)____

   2)____

5. How much confidence do you place in your partner's ability to accomplish this mission?

1	2	3	4	5	6	7
Very Low			Moderate			Very High

Explain briefly:____

6. How much confidence do you think your partner places in your ability to accomplish this mission?

1	2	3	4	5	6	7
Very Low			Moderate			Very High

Explain briefly:____

CREW-MEMBER POST-MISSION QUESTIONNAIRE

CREW ID: ________________	CREW POSITION: Pilot / CPG
DATE/TIME: ________________	LAST 4 DIGITS OF SS#: ________

1. How much confidence did you have in your fellow crew member, as compared to flying with other aviators in this unit?

1	2	3	4	5	6	7
Much Less			About The Same			Much More

2. How much assistance did you provide your fellow crew member, as compared to flying with other aviators in this unit?

1	2	3	4	5	6	7
Much Less			About The Same			Much More

3. How much did you cross-monitor the actions of your fellow crew member, as compared to flying with other aviators in this unit?

1	2	3	4	5	6	7
Much Less			About The Same			Much More

4. To what extent were you able to anticipate (i.e., predict) the actions and decisions of your fellow crew member?

1	2	3	4	5	6	7
Rarely			Half The Time			All The Time

5a. What was the most critical episode of this mission?______________________________

____

b. During this critical episode to what extent were you thinking and acting "in sync" with your fellow crew member?

1	2	3	4	5	6	7
Not At All			Moderately			A Great Deal

c. How do you know that? ____

____

6. What is the one thing you would do differently if you flew the same mission again?

____

IP POST-MISSION TEAMWORK RATING FORM

CREW ID:________________________ IP:__________________

DATE/TIME:________________________

INSTRUCTIONS

Circle a number on the scale accompanying the questions below so that it best describes the behavior of the crew you just observed. Try to rate the behavior of the crew on an absolute scale and not a relative scale that compares one crew to the next. To help you perform this absolute rating a brief description of the behavior you should observe for the highest rating on the scale and a brief description of the behavior you should observe for the lowest rating on the scale are provided for each question. Read these guides or anchors carefully and refer to them as you rate the crew on each question. Feel free to write comments or explanations to any question.

Team Orientation

1. To what extent was this crew oriented toward teamwork?

1	2	3	4	5	6	7

7 Good team orientation could be inferred in a situation where a team member places the goals and interests of the team ahead of personal goals. Also may be evident through the display of trust, team pride, and esprit de corps, and an awareness that teamwork is important.

1 Poor team orientation manifests itself when members place their personal concerns above the team's success (e.g., disregarding or refusing to follow procedures; arguments, quarrels, and open resentment; and becoming upset with a member's performance and either ignoring or harassing that member are evidences of poor team orientation).

2. To what extent were errors caused by inadequate crew communication?

1	2	3	4	5	6	7

7 Communication within the crew was always effective and never responsible for errors or degraded performance.

1 Communication was wholly inadequate and resulted in most of the errors made by the crew.

3. To what extent were errors caused by inadequate individual actions?

1	2	3	4	5	6	7

7 No actions of a single crew member resulted in errors or poor crew performance

1 The actions and decisions by a single crew member very frequently resulted in errors or poor crew performance.

Comments:____

Communication Behavior

4. How well did crew members communicate?

1	2	3	4	5	6	7

7 Good communication occurs when team members pass on all important information and clarify intentions and planned procedures; members obtain necessary information and acknowledge and repeat messages to ensure correctness; members ensure that their messages are received as intended.

1 Poor communication occurs when crew members fail to pass on information or intentions, or pass on incomplete communications; members fail to clarify information; members fail to acknowledge other member's requests or reports; members disregard proper security procedures for communication; members use improper terminology; members tie up the net with irrelevant communications.

5. To what extent did the pilot provide relevant information to the CPG, without the CPG having to ask for it?

1	2	3	4	5	6	7

7 Pilot always provided important information to the CPG without being asked.

1 Pilot never provided information to the CPG unless specifically asked.

6. To what extent did the CPG provide relevant information to the pilot, without the pilot having to ask for it?

1	2	3	4	5	6	7

7 CPG always provided important information to the pilot without being asked.

1 CPG never provided information to the pilot unless specifically asked.

Comments:____

Monitoring Behavior

7. To what extent did crew members monitor each other's behavior?

1	2	3	4	5	6	7

7 Good monitoring occurs when crew members consistently observe the performance of the other to ensure the efficiency of the team; members notice and are concerned with the performance of the entire team; one member recognizes when other crew member performs correctly; consistently keeps track of other crew member's performance.

1 Poor monitoring occurs when one crew member fails to notice the other's performance on almost all occasions; rarely notices when the other crew member performs correctly or makes a mistake.

8. To what extent did crew members alert each other to impending decisions and actions?

1	2	3	4	5	6	7

7 Crew members always alerted each other to impending decisions and actions; supporting information was actively solicited from other crew member.

1 Crew members did not keep each other informed of impending decisions and actions; compromises to flight safety or mission effectiveness arose when a crew member waited for the other to volunteer significant information.

Comments:____

Feedback Behavior

9. To what extent did crew members provide feedback to each other?

1	2	3	4	5	6	7

7 Good feedback behavior occurs when crew members go over procedures with each other by identifying mistakes and how to correct them; ask for input regarding mistakes and what needs to be worked on; members are corrected for mistakes and incorporate the suggestions in their procedures.

1 Poor feedback behavior occurs when one crew member makes sarcastic comments to other when the scenario doesn't go as planned; resists asking for advice and makes guesses on proper procedures; rejects time-saving suggestions offered by other crew member.

Comments:____

Backup Behavior

10. To what extent did crew members provided backup to each other?

1	2	3	4	5	6	7

7 Good backup behavior occurs when one crew member is having difficulty, makes a mistake, or is unable to perform duties, and the other member steps in to help ensuring that the activity is completed properly; one member provides critical assistance without neglecting their own assigned duties; member having difficulty or is overburdened displays a willingness to seek assistance rather than struggle and make a mistake.

1 Poor backup behavior occurs when one crew member fails to provide assistance to other member who is having difficulty, makes a mistake, or is unable to perform his duties; while providing assistance, the member tends to neglect his own duties; member is unwilling to ask for help even when it is available; one member provides needed assistance, but does not inform other that they are assisting of what they he has done; one member displays an unwillingness to help other even when asked.

11. To what extent did the pilot anticipate the need to provide task assistance to the CPG?

1	2	3	4	5	6	7

7 Pilot consistently anticipated the need to provide task assistance to CPG during critical phases of flight.

1 Pilot never anticipated the need to provide task assistance to CPG during critical phases of flight; the CPG always had to ask.

12. To what extend did the CPG anticipate the need to provide task assistance to the pilot?

1	2	3	4	5	6	7

7 CPG consistently anticipated the need to provide task assistance to pilot during critical phases of flight.

1 CPG never anticipated the need to provide task assistance to pilot during critical phases of flight; the pilot always had to ask.

13. Did the crew members adjust individual task responsibilities to prevent overload?

1	2	3	4	5	6	7

7 Crew members were consistently aware of workload buildup on each others and reacted quickly to adjust division of task responsibilities to redistribute workload among each other.

1 Crew members were generally unaware of workload buildup on each others; little or no attempt was made to adjust the distribution of task responsibilities before significant compromises to flight safety or mission effectiveness occur.

Comments:____

Coordination Behavior

14. To what extent was the crew's behavior coordinated?

1	2	3	4	5	6	7

7 Good coordination behavior occurs when one team member consistently passes critical information to the other member, thereby enabling him/her to accomplish tasks; one member consistently carries out tasks quickly or in a timely manner enabling other to carry out his tasks effectively. Crew members appear very familiar with the relevant parts of each other's job and carry out individual tasks in a synchronized manner.

1 Poor coordination behavior occurs when one team member consistently carries out his tasks ineffectively, leading to other team member failing at his tasks; members carry out their tasks unpredictably, leading to delays in execution of critical tasks; members neglect to pass on critical pieces of information to each other, leading to breakdown in team performance; team members carry out their tasks with significant delays leading to crew errors.

15. How congruent/similar were the pilot's and CPG's understandings of the mission?

1	2	3	4	5	6	7

7 Pilot and CPG were completely in agreement (i.e., congruent) on all goals, tasks, and concepts involving the mission.

1 Pilot and CPG were rarely in agreement (i.e., congruent) on all goals, tasks, and concepts involving the mission.

Comments:____

Summary

How would you characterize the most salient aspect of this crew's teamwork skills(positive and negative)?

What recommendation would you make to this crew to improve their crew coordination and teamwork skills?

Appendix B

Team Coordination Evaluation Instruments

TEAMWORK DIMENSION RATING FORM

Tape # __________ Observer: ________________

1. Team Orientation:

1	2	3	4	5	6	7

7 Positive Team Orientation could be inferred in a situation where a team member places the goals and interests of the team ahead of personal goals. Also may be evident through the display of trust, team pride, and esprit de corps. Examples: members...

- willingly participate in all relevant aspects of the team
- cooperate fully with one another
- display a high degree of pride in their duties and the team
- display an awareness that they are part of a team & that teamwork is important

1 Poor Team Orientation manifests itself when members place their personal concerns above the team's success (e.g., disregarding or refusing to follow procedures; arguments, quarrels, and open resentment; and becoming upset with a member's performance and either ignoring or harassing that member are evidences of poor team orientation). Examples: members...

- display a high degree of distrust for one another
- allow interpersonal conflicts to interfere with team functioning
- place their own personal goals and interests ahead of the goals of the team

2. Communication Behavior:

1	2	3	4	5	6	7

7 Positive Communication occurs when team members pass on all important information and clarify intentions and planned procedures; members obtain necessary information and acknowledge and repeat messages to ensure correctness; members ensure that their messages are received as intended. Examples: members...

- pass complete information as prescribed
- follow proper procedures and terminology in passing and receiving information
- acknowledge requests of each other

1 Poor Communication occurs when members fail to pass on information or intentions, or pass on incomplete communications; members fail to clarify information; members fail to acknowledge other member's requests or reports; members disregard proper security procedures for communication; members use improper terminology; members tie up the net with irrelevant communications. Examples: members...

- fail to clarify intentions to each other
- tie up the net with excessive chatter
- do not ensure that recipient of information understood

3. Monitoring Behavior:

1	2	3	4	5	6	7

7 Positive Monitoring occurs when team members consistently observe the performance of the other to ensure the efficiency of the team; members notice and are concerned with the performance of the entire team; one member recognize when other team member performs correctly; consistently keep track of other team member's performance. Examples: one team member ...

- is aware of other team member's performance
- recognizes when other team member makes a mistake or performs correctly
- listens to other team member to make sure other is performing correctly

1 Poor Monitoring occurs when one team member fails to notice the others performance on almost all occasions; rarely notices when other team member performs correctly or makes a mistake. Examples: one team member ...

- is not concerned with the performance of other team member
- does not ensure that other team member is performing appropriately

4. Feedback Behavior:

1	2	3	4	5	6	7

7 Positive Feedback Behavior occurs when team members go over procedures with one another by identifying mistakes and how to correct them; ask for input regarding mistakes and what needs to be worked on; members are corrected for mistakes and incorporate the suggestions in their procedures. Examples: members...

- respond to each others requests for performance information, if necessary
- ask for advice on proper procedures
- are corrected on a few mistakes, and incorporate the suggestions into their procedures

1 Poor Feedback Behavior occurs when one team member makes sarcastic comments to other when the scenario doesn't go as planned; resists asking for advice and makes guesses on proper procedures; rejects time-saving suggestions offered by other team member. Examples: members...

- provide unneeded suggestions that ultimately confuse
- inform each other that all is doing great, rather than giving specific and constructive advice

5. Backup Behavior:

1	2	3	4	5	6	7

7 Positive Backup Behavior occurs when one team member is having difficulty, makes a mistake, or is unable to perform duties, and the other member steps in to help ensuring that the activity is completed properly; one member provides critical assistance without neglecting their own assigned duties; member having difficulty or is overburdened displays a willingness to seek assistance rather than struggle and make a mistake. Examples: members...

- ask for help when needed rather than struggle
- provide assistance to each other when difficulty is observed, even when not asked
- ensure that each is aware of what was done to help

1 Poor Backup Behavior occurs when one team member fails to provide assistance to other member who is having difficulty, makes a mistake, or is unable to perform his duties; while providing assistance, the member tends to neglect his own duties; member is unwilling to ask for help even when it is available; one member provides needed assistance, but does not inform other that they are assisting of what they he has done; one member displays an unwillingness to help other even when asked. Examples: members...

- fail to provide assistance to each other when specifically asked
- neglect their own duties in the process of helping
- fail to provide assistance to each other when it is obviously needed

6. Coordination Behavior:

1	2	3	4	5	6	7

7 Positive Coordination Behavior occurs when one team member consistently passes tactical information to other member, thereby enabling him to accomplish tasks; one member consistently carries out tasks quickly or in a timely manner enabling other to carry out his tasks effectively. Examples: members...

- pass performance-relevant data from one to the other in a timely fashion when required
- are familiar with the relevant parts of each other's job

• carries out individual tasks in a synchronized manner

1 Poor Coordination Behavior occurs when one team member consistently carries out his tasks ineffectively, leading to other team member failing at his tasks; members carry out their tasks unpredictably, leading to delays in execution of critical tasks; members neglect to pass on critical pieces of information to each other, leading to breakdown in team performance; team members carry out their tasks with significant delays leading to the team's failure. Exempts: members...

- pass performance-relevant data from one to the other in an inefficient manner
- hinder the performance of each other
- cause distractions during critical operations

RATER'S SUBJECTIVE EVALUATION FORM

OBSERVATION OF CAMPBELL TAPES

Tape # ________ Observer: ______________

1. Did the pilot anticipate the need to provide information and warnings to the CPG?

1	2	3	4	5	6	7

7 Pilot consistently anticipated the need to provide information or warnings to the CPG during critical phases of flight

4 Pilot usually anticipated the need to provide information or warnings to the CPG during critical phases of flight, but sometimes the CPG had to ask

1 Pilot never anticipated the need to provide information or warnings to the CPG during critical phases of flight; the CPG always had to ask

2. Did the pilot anticipate the need to provide task assistance to the CPG?

1	2	3	4	5	6	7

7 Pilot consistently anticipated the need to provide task assistance to CPG during critical phases of flight

4 Pilot usually anticipated the need to provide task assistance to CPG during critical phases of flight, but sometimes the CPG had to ask

1 Pilot never anticipated the need to provide task assistance to CPG during critical phases of flight; the CPG always had to ask

3. Did the CPG anticipate the need to provide information and warnings to the pilot?

1	2	3	4	5	6	7

7 CPG consistently anticipated the need to provide information or warnings to the pilot during critical phases of flight

4 CPG usually anticipated the need to provide information or warnings to the pilot during critical phases of flight, but sometimes the pilot had to ask

1 CPG never anticipated the need to provide information or warnings to the pilot during critical phases of flight; the pilot always had to ask

4. Did CPG anticipate the need to provide task assistance to the pilot?

1	2	3	4	5	6	7

7 CPG consistently anticipated the need to provide task assistance to pilot during critical phases of flight

4 CPG usually anticipated the need to provide task assistance to pilot during critical phases of flight, but sometimes the pilot had to ask

1 CPG never anticipated the need to provide task assistance to pilot during critical phases of flight; the pilot always had to ask

5. To what extent did the pilot appear to understand (i.e., know the reason for) the actions and decisions of the CPG?

1	2	3	4	5	6	7

7 Pilot always understood and knew the reasons for all the CPG's actions and decisions

4 Pilot understood and knew the reasons for the CPG's actions and decisions about half the time

1 Pilot never understood and knew the reasons for the CPG's actions and decisions

6. To what extent did the CPG appear to understand (i.e., know the reason for) the actions and decisions of the pilot?

1	2	3	4	5	6	7

7 CPG always understood and knew the reasons for all the pilot's actions and decisions

4 CPG understood and knew the reasons for the pilot's actions and decisions about half the time

1 CPG never understood and knew the reasons for the pilot's actions and decisions

7. How were conflicts handled between crew members?

1	2	3	4	5	6	7

7 Disagreements were handled in a professional manner without involving personal attacks or defensive posturing

4 Disagreements did not involve obvious attacks of character or defensive posturing

1 Conflicts involved personal attacks and resulted in a disruption of teamwork

8. Did the crew take advantage of low workload periods to rehearse up-coming flight segments?

1	2	3	4	5	6	7

7 Crew members took advantage of low workload periods during mission to rehearse upcoming flight segments

4 Crew members engaged in some in-flight rehearsal of up-coming flight segments; no major coordination problems arose that could be attributed to a failure to rehearse

1 Little or no attention was given to in-flight rehearsal of up-coming flight segments; some coordination problems were attributed to a failure to rehearse

9. Did the crew members adjust individual task responsibilities to prevent overload?

1	2	3	4	5	6	7

7 Crew members were consistently aware of workload buildup on each others and reacted quickly to adjust distribution of task responsibilities

4 Crew members maintained some awareness of workload buildup on each others; workload was adjusted before serious compromise to flight safety or mission effectiveness occurs

1 Crew members were generally unaware of workload buildup on each others; little or no attempt was made to adjust the distribution of task responsibilities before significant compromises to flight safety or mission effectiveness occur

10. Did crew members keep each other aware of critical mission factors, obstacles, or conditions?

1	2	3	4	5	6	7

7 Crew members routinely updated each other on critical mission factors, obstacles, or conditions; significant changes were highlighted and acknowledged

4 Crew members occasionally updated each other on critical mission factors, obstacles, or conditions; no significant compromises to flight safety or mission effectiveness occurred

1 Crew members disregarded need to keep each other informed of critical mission factors, obstacles, or conditions; significant compromises to fight safety or mission effectiveness occurred as a result

11. Did the crew members alert each others to impending decisions and actions?

1	2	3	4	5	6	7

7 Crew members alerted each other to impending decisions and actions; supporting information was actively solicited from other crew member

4 Crew members occasionally solicited information from each other regarding impending decisions and actions; no significant compromises to flight safety or mission effectiveness occurred

1 Crew members did not keep each other informed of impending decisions and actions; compromises to flight safety or mission effectiveness arose when a crew member waited for others to volunteer significant information

12. Did crew members justify their recommended plans and courses of action with appropriate rationale?

1	2	3	4	5	6	7

7 Time permitting, crew members consistently provided rationale for their recommended plans and courses of action; a professional atmosphere was maintained

4 When misunderstanding was apparent, crew members provided rationale for their recommended plans and courses of action; some level of objectivity was maintained

1 Crew members frequently justified their recommendations on rank or experience level, rather than logic; personality conflicts resulted from this behavior

Appendix C

Analysis of Behavioral, Performance, and Attitudinal Measures from Aircrew Coordination Testbed

Table C-1

Explanation of Mission Performance Measures

Table C.1	Description	Component Measures
PERF1	Terrain Flight Navigation	Total deviations in Segment 1 + Total deviations in Segment 2
PERF2	Threat Avoidance and Evasion	Total deviations in Segment 1 + Total deviations in Segment 2
PERF3	Aircraft Emergencies	Diagnosis of major emergency + Eventual outcome
PERF4	Unexpected Event	Eventual outcome of IMC
PERF5	Instrument Flight Recovery	Rating of approach planning × Approach execution outcome
PERF6	Mission Threatening Crew Error	Mission error rating
% Subobj	Percent of Subobjectives Completed	Percent of 13 subobjectives completed
Total PERF	Overall Performance Measure	PERF1 + PERF2 + PERF3 + PERF4 + PERF5 + PERF6

Table C-2

Differences between Pre- and Post-Coordination Training on Behavioral, Performance, and Attitudinal Measures

Measure	Pre-Training Mean	Post-Training Mean	Difference	significance (p)
Overall Grade	0.75	1.81	1.06	.004
ACE Basic Qualities (BQ)				
Average ACE	3.17	4.30	1.13	.001
BQ1	3.94	5.06	1.13	.006
BQ2	3.06	4.63	1.56	.003
BQ3	3.19	4.31	1.13	.001
BQ4	2.94	4.44	1.50	.001
BQ5	2.94	4.13	1.19	.001
BQ6	3.13	4.44	1.31	.003
BQ7	2.81	3.75	0.94	.05
BQ8	3.25	4.56	1.31	.006
BQ9	3.25	4.19	0.94	.005
BQ10	2.69	3.88	1.19	.006
BQ11	3.19	4.19	1.00	.02
BQ12	3.19	4.13	0.94	.003
BQ13	3.69	4.19	0.50	ns

Table C-2 (continued)

Differences between Pre- and Post-Coordination Training on Behavioral, Performance, and Attitudinal Measures

Measure	Pre-Training Mean	Post-Training Mean	Difference	significance (p)
ATM Measures				
Average ATM	1.36	1.93	0.57	.001
1000 Mission brief	0.93	2.13	1.20	.001
1004 PPC	1.73	2.07	0.34	ns
1007 Start/run up	1.67	2.00	0.33	.10
1016 Hover power check	0.33	1.80	1.57	.001
1018 VMC takeoff	2.00	2.20	0.20	.08
1023 Fuel mgt procedure	0.27	1.27	1.00	.008
1026 Elect-aided nav (** not reported **)				
1028 VMC approach	1.87	2.00	0.13	ns
1068 Emergency	1.13	1.60	0.47	ns
1076 Radio navigation	1.80	2.13	0.33	ns
1081 Non.precision approach	1.20	1.73	0.53	ns
1083 VHIRP	1.87	2.20	0.33	ns
1095 Operate ASE	1.53	2.00	0.47	.03
2008 Evasive maneuvers	1.47	1.87	0.40	ns
2009 Multiaircraft operations	1.53	1.87	0.44	.10
2016 Expernal load operations	1.20	2.00	0.80	.003
2078 Terrain flight msn plan	1.27	2.27	1.00	.001
2079 Terrain flight navigation	1.00	2.00	1.00	.001
2081 Terrain flight	1.80	2.00	0.20	ns
Mission Performance Measures				
Total Performance	20.1	19.8	-0.30	ns
PERF1	4.56	3.13	-1.43	.033
PERF2	5.67	5.44	0.23	ns
PERF3	3.75	3.69	-0.06	ns
PERF4	1.88	2.00	0.12	.164
PERF5	3.19	3.56	0.37	ns
PERF6	1.06	1.94	0.88	.011
Percent subobjectives completed	66.3	85.8	19.5	.004
CMAQ Measures				
Average CMAQ	5.56	5.71	0.15	.003
CMAQ Index 1	4.04	4.18	0.14	.057
CMAQ Index 2	5.87	6.10	0.23	.071

Table C-3

Correlations Between Behavioral Evaluation Measures and Overall Grade Before and After Coordination Training Course

	Pre-Coordination Training		Post-Coordination Training	
Measure	Correlation	significance (p)	Correlation	significance (p)
Average ATM	0.73	.002	0.73	.002
Average ACE	0.67	.003	0.74	.001
BQ1	0.43	.05	0.47	.04
BQ2	0.66	.004	0.66	.004
BQ3	0.31	ns	0.58	.01
BQ4	0.66	.004	0.62	.006
BQ5	0.47	.04	0.43	.05
BQ6	0.61	.007	0.66	.004
BQ7	0.55	.02	0.69	.002
BQ8	0.50	.03	0.66	.004
BQ9	0.23	ns	0.69	.002
BQ10	0.43	.05	0.65	.004
BQ11	0.33	ns	0.69	.002
BQ12	0.23	ns	0.34	ns
BQ13	0.15	ns	0.35	ns

Appendix D

Correlations of Communications Measures with ATM-Based Measures

Table D-1

Correlations of Communications Measures with ATM Measures Before Training

	ATM Tasks Pre-Training (Part 1 of 2)									
Communication Measures	1000 Mission brief	1004 PPC	1007 Start/ Run up	1016 Hover pwr ck	1018 VMC takeoff ***	1023 Fuel mgt proced	1028 VMC appr	1068 Emer-gency	1076 Radio nav	1081 Non-precis. appr
ANTICIPATION RATIOS										
Total Anticipation Ratio	.07	-.43*	.51*	-.20		.07	.00	.31	-.01	.15
Anticipation Ratio for Information Transmissions	.40‡	-.33	.44*	-.34		.28	-.03	-.08	-.06	-.03
Anticipation Ratio for Actions and Tasks	.23	-.37‡	.33	-.33		-.01	.05	.27	-.08	-.09
Anticipation Ratio in Routine Conditions	.12	.45*	.45*	-.26		.01	.01	.23	-.05	.11
Anticipation Ratio in Crisis Conditions	-.19	-.26	.49*	-.37‡		.31	.04	.47*	.13	.22
COMMUNICATION RATES										
Communication Rate—Total	.33	-.11	.12	.38‡		-.36‡	-.16	.02	-.12	-.04
Proportion of Communication Rate for Information Transmissions	.10	-.01	.22	.49*		.09	.44*	.13	.27	.27
Proportion of Communication Rate for Actions and Tasks	-.11	.10	-.21	-.39‡		.07	-.18	-.33	-.32	-.73**
Proportion of Communication Rate for Planning and Problem Solving	-.01	-.11	-.05	-.20		-.22	-.40‡	.23	.01	.52*
Comm. Rate in Routine Conditions	.27	-.09	.06	.34		-.36‡	-.16	-.01	-.16	-.09
Comm. Rate in Crisis Conditions	.59*	-.18	.42‡	.50*		-.29	-.03	.24	.14	.25
Proportion of Communication Rate for Information Transmissions in Routine Conditions	.17	.02	.20	.56*		.11	-.45*	.17	.24	.23
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	-.17	.07	-.21	-.46*		.06	-.24	-.35‡	-.37‡	-.69**
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	-.03	-.13	.00	-.20		-.23	-.34	.29	.14	.58*
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	-.29	-.19	.27	-.12		.03	.14	.13	.26	.43‡
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	.27	.16	-.08	.17		.00	.26	.02	.15	-.53*
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	.06	.06	-.24	-.05		-.05	-.46*	-.19	-.50*	.06
Number of Significant Correlations	2	3	5	7		2	4	2	2	6

‡ $p \leq .10$

* $p \leq .05$

** $p \leq .01$

*** No Variance

Table D-1 (continued)

Correlations of Communications Measures with ATM Measures Before Training

	ATM Tasks Pre-Training (Part 2 of 2)									
Communication Measures	1083 VHIRP	1095 Operate ASE	2008 Evasive maneuvers	2009 Multi-aircraft opns	2016 External load opns	2078 Terrain flt msn plan	2079 Terrain flt nav	2081 Terrain flt	Average ATM Score	Number of signif correl.
ANTICIPATION RATIOS										
Total Anticipation Ratio	.02	.30	-.01	.22	.24	.13	.50*	-.09	.25	3
Anticipation Ratio for Information Transmissions	-.06	.24	.08	-.45*	.36‡	-.17	.15	.07	-.01	4
Anticipation Ratio for Actions and Tasks	.00	.14	-.07	.49*	.05	.02	.38‡	-.26	.11	3
Anticipation Ratio in Routine Conditions	-.05	.27	.02	.26	.18	.16	.49*	-.19	.19	3
Anticipation Ratio in Crisis Conditions	.18	.32	.00	-.02	.45*	-.09	.57*	.22	.37‡	6
COMMUNICATION RATES										
Communication Rate—Total	-.50*	-.42‡	.27	.53*	-.32	.15	.17	.07	-.06	5
Proportion of Communication Rate for Information Transmissions	.02	.23	.15	-.18	.56*	.62**	.03	-.07	.42‡	5
Proportion of Communication Rate for Actions and Tasks	-.07	-.38‡	-.15	.09	-.48	-.66**	-.37‡	-.03	-.57*	7
Proportion of Communication Rate for Planning and Problem Solving	.07	.14	-.02	.14	-.21	-.06	.40	.14	.12	3
Comm. Rate in Routine Conditions	-.48*	-.43‡	.26	.52*	-.36‡	.08	.15	.04	-.11	5
Comm. Rate in Crisis Conditions	-.36‡	-.19	.22	.47*	.08	.57‡	.18	.17	.33	6
Proportion of Communication Rate for Information Transmissions in Routine Conditions	-.03	.16	.21	-.10	.50*	.63**	.03	-.05	.41‡	5
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	-.03	-.34	-.18	.07	-.49*	-.74**	-.36‡	-.10	-.61**	8
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	.08	.22	-.07	.05	-.08	.08	.44*	.22	.20	2
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	.23	.50*	-.18	-.50*	.57*	.26	-.07	-.10	.27	4
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	-.24	-.39‡	.03	.24	-.10	.18	-.11	.29	-.06	2
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	.02	-.19	.20	.37‡	-.61**	-.53*	.04	-.20	-.27	5
Number of Significant Correlations	3	5	0	7	9	6	8	0	5	76

‡ $p \leq .10$

* $p \leq .05$

** $p \leq .01$

*** No Variance

Table D-2

Correlations of Communications Measures with ATM Measures After Training

Communication Measures	ATM Tasks Post-Training (Part 1 of 2)									
	1000 Mission brief	1004 PPC	1007 Start/ Run up	1016 Hover pwr ck	1018 VMC takeoff	1023 Fuel mgt proced	1028 VMC appr	1068 Emer-gency	1076 Radio nav	1081 Non-precis. appr
ANTICIPATION RATIOS										
Total Anticipation Ratio	.25	.58*	.76**	.12	.43‡	.28	.59*	-.11	-.37‡	.09
Anticipation Ratio for Information Transmissions	.09	.42‡	.54*	-.09	.14	.35‡	.34	-.22	-.34	-.10
Anticipation Ratio for Actions and Tasks	.21	.40‡	.55*	.11	.17	.04	.41	.00	-.31	.02
Anticipation Ratio in Routine Conditions	.26	.54*	.75**	.12	.41‡	.26	.55*	-.13	-.33	.13
Anticipation Ratio in Crisis Conditions	.04	.51*	.41‡	.13	.30	.27	.42‡	.05	-.44*	-.11
COMMUNICATION RATES										
Communication Rate—Total	.35‡	-.01	.07	.07	-.12	.42‡	-.08	.16	.45*	.25
Proportion of Communication Rate for Information Transmissions	.00	.01	.14	-.06	-.09	-.06	-.07	-.48*	-.26	-.21
Proportion of Communication Rate for Actions and Tasks	-.03	-.11	-.22	-.17	-.32	-.09	-.17	-.25	.15	-.12
Proportion of Communication Rate for Planning and Problem Solving	.03	.11	.07	.27	.49*	.14	.30	-.35‡	.17	.41‡
Comm. Rate in Routine Conditions	.30	.01	.05	.07	-.16	.43‡	-.11	.16	.38‡	.21
Comm. Rate in Crisis Conditions	.51*	-.18	.13	.14	.12	.34	.07	.18	.73**	.42‡
Proportion of Communication Rate for Information Transmissions in Routine Conditions	.04	.10	.23	-.09	-.06	.02	-.02	.41‡	-.25	-.12
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	-.04	-.18	-.30	-.11	-.30	-.18	-.19	-.21	.15	-.15
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	.01	.08	.06	.27	.47*	.20	.28	-.31	.16	.36‡
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	-.12	-.32	-.24	.14	-.15	-.22	-.28	.52	-.18	-.41‡
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	.06	.24	.21	-.40‡	-.25	.23	.04	-.30	-.08	.03
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	.11	.20	.11	.26	.51*	.06	.36	-.42‡	.17	.57*
Number of Significant Correlations	2	5	5	1	5	3	4	5	5	5

‡ $p \leq .10$

* $p \leq .05$

** $p \leq .01$

*** No Variance

Table D-2 (continued)

Correlations of Communications Measures with ATM Measures After Training

Communication Measures	ATM Tasks Post-Training (Part 2 of 2)									
	1083 VHIRP	1095 Operate ASE	2008 Evasive maneu-vers	2009 Multi-aircraft opns	2016 Exter-nal load opns	2078 Terrain flt msn plan	2079 Terrain flt nav	2081 Terrain flt	Avera-ge ATM Score	Num-ber of signif correl.
ANTICIPATION RATIOS										
Total Anticipation Ratio	.20	.08	.02	.67**	.25	.17	.18	.04	.41‡	7
Anticipation Ratio for Information Transmissions	.12	.09	-.13	.28	.16	-.07	.04	-.07	.13	3
Anticipation Ratio for Actions and Tasks	.21	.24	.09	.56*	.07	.06	.14	-.25	.28	4
Anticipation Ratio in Routine Conditions	.21	.14	-.03	.70**	.17	.13	.08	-.02	.38‡	6
Anticipation Ratio in Crisis Conditions	-.11	-.20	.30	.26	.35‡	.11	.54*	.29	.34	6
COMMUNICATION RATES										
Communication Rate—Total	.25	.10	-.32	.19	-.04	.09	-.03	-.27	.17	3
Proportion of Communication Rate for Information Transmissions	.02	-.21	.31	.15	-.01	-.08	.08	.09	.07	1
Proportion of Communication Rate for Actions and Tasks	.05	.36‡	-.29	-.25	-.12	-.12	-.15	-.50*	-.26	2
Proportion of Communication Rate for Planning and Problem Solving	-.08	-.14	-.07	.09	.16	.26	.06	.47*	.22	4
Comm. Rate in Routine Conditions	.17	.08	-.30	.18	-.05	.03	-.02	-.28	.14	2
Comm. Rate in Crisis Conditions	.51*	.19	-.30	.20	.00	.33	.03	-.18	.42‡	5
Proportion of Communication Rate for Information Transmissions in Routine Conditions	.04	-.18	.24	.26	-.02	-.10	.04	.16	.09	1
Proportion of Communication Rate for Actions and Tasks in Routine Conditions	.02	.33	-.23	-.33	-.10	.09	-.10	-.55*	-.27	1
Proportion of Communication Rate for Planning and Problem Solving in Routine Conditions	-.08	-.16	-.05	.05	.15	.25	.06	.47*	.20	3
Proportion of Communication Rate for Information Transmissions in Crisis Conditions	-.16	-.22	.35‡	-.24	.02	-.09	.15	-.29	-.10	3
Proportion of Communication Rate for Actions and Tasks in Crisis Conditions	.28	.34	-.35‡	.14	-.16	-.13	-.25	-.01	-.07	2
Proportion of Communication Rate for Planning and Problem Solving in Crisis Conditions	.08	-.07	-.11	.19	.15	.29	.07	.43‡	.24	4
Number of Significant Correlations	1	1	2	3	1	0	1	5	3	57

‡ $p \leq .10$

* $p \leq .05$

** $p \leq .01$

*** No Variance